ESCAP MEETING NO. 50 - 06/18/01

AGENDA

Kathleen P Porter
06/15/2001 02:24 PM

To: Angela Frazier/DMD/HQ/BOC@BOC, Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol M Van Horn/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Vanessa M Leuthold/DMD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC
cc:
Subject: ESCAP Agenda for June 18

The agenda for the June 18 ESCAP Meeting scheduled from 10:30-12 in Rm. 2412/3 is as follows:

1. ESCAP process for reviewing data

2. Technical Issues

3. Issuance of the research plan

4. Potential meeting of outside experts

ESCAP MEETING NO. 50 - 06/18/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 50

June 18, 2001

Prepared by: Nick Birnbaum

The fiftieth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on June 18, 2001 at 10:30 am. The agenda for the meeting was to discuss: 1) the Committee's schedule for reviewing data and analyses (to inform their recommendations regarding the potential use of the adjusted data for purposes other than redistricting), 2) over-arching technical issues, 3) the issuance of a research plan, and 4) the possibility of conducting meetings with external experts before recommendations are issued.

Committee Attendees:

Ruth Ann Killion
Cynthia Clark
John Thompson
Howard Hogan
Carol Van Horn
Bob Fay
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines
Fay Nash
Nick Birnbaum
Kathleen Styles
Bill Bell
Raj Singh
Rita Petroni
Sarah Brady
Donna Kostanich
Tommy Wright

I. Schedule for Reviewing Data and Analyses

John Thompson briefly discussed some procedural matters with regard to the Committee's upcoming work. Data and analyses will be available as soon as the end of June. It is expected that all the data that the Committee plans to examine in reaching its recommendations regarding the adjusted data will be available no later than the end of September. Beginning with this meeting, the Committee will now meet weekly or more frequently. Meetings have been scheduled through mid-October. Despite recent changes within the senior staff of the Census

Bureau, it was noted that John Thompson would continue to chair these meetings.

II. Technical Issues for the Committee's Consideration

1) Re-defining the post-strata - There was some discussion about whether the Committee should contemplate redefining the post-strata. It was argued that, given the time frame for making recommendations regarding the potential use of the adjusted data for purposes other than redistricting, and legitimate questions about whether a different post-stratification scheme would result in more accurate estimates, this did not appear to be a course of action the Committee should pursue. It was noted that future research could consider new information in defining post-strata, but that research would be conducted well after this next decision.

2) Duplication - There was also discussion about evaluations directed at uncovering the level of duplication in Census 2000 that was outside the scope of the A.C.E. One issue discussed was whether to use the results in combination with the A.C.E. to produce a revised adjustment. It was determined that (1) it would be important to consider the results of the duplication evaluations in conjunction with the A.C.E.; and (2) the Committee would examine the level of any duplication to determine whether any revised adjustments should be calculated.

3) Re-Calculating the Adjustment Estimates - The discussion of the above-mentioned items segued into a more general discussion about under what circumstances the Committee would pursue re-calculating the adjustment estimates based on information learned from the evaluation data and other studies currently being conducted. What are the technical criteria for this decision? Can one define numeric thresholds with regard to, for example, P and E sample errors, matching errors, etc., to determine it is appropriate to produce a revised set of DSEs? It was noted that there would be variance/bias tradeoffs with regard to producing revised adjustment numbers; that is, reductions in bias would be offset (at least somewhat) by increases in variance, given the sizes of the samples from which the evaluation data are obtained. John Thompson requested that Howard Hogan's staff examine these variance/bias tradeoffs to develop appropriate criteria for such a decision.

III. Issuance of the Committee's Research Plan and Potential Meetings with External Experts

John Thompson requested that Committee members provide comments on the initial rough draft research plan by the end of the week.

Next, the Committee discussed the possibility of conducting discussions with external stakeholders once all the data have been analyzed and presented to the Committee. The discussion touched on the possible composition of such an external group – for example, federal agency representatives and other experts – and the method(s) for obtaining their input (public meetings, hiring individuals as consultants, etc.). It was clear that the Committee needed to determine if the evaluation results would raise issues that external input could help to address.

It was decided that the Committee would continue to discuss this issue at future meetings.

IV. Next Meeting

The next meeting is scheduled for June 25, 2001. The agenda for that meeting is to discuss the planned research on Demographic Analysis and the consultation of external experts on this issue.

ESCAP MEETING NO. 51 - 06/25/01

AGENDA

Kathleen P Porter
06/25/2001 08:12 AM

To: Angela Frazier/DMD/HQ/BOC@BOC, Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol M Van Horn/DMD/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DMD/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Vanessa M Leuthold/DMD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC
cc:
Subject: Agenda for 6/25 ESCAP meeting

The agenda for today's ESCAP meeting in Rm. 2412/3 at 10:30:

Comparison of Population Change: External Consultation and Current Research - John Long

ESCAP MEETING NO. 51 - 06/25/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 51

June 25, 2001

Prepared by: Sarah Brady

The fifty-first meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on June 25, 2001 at 10:30. The agenda for the meeting was to discuss the external consultation and current research underway for demographic analysis.

Committee Attendees:

Nancy Potok
John Thompson
Cynthia Clark
Nancy Gordon
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
John Long

Other Attendees:

Marvin Raines	Raj Singh
Bill Bell	Donna Kostanich
Tommy Wright	Rita Petroni
Gregg Robinson	Fay Nash
Signe Wetrogan	Maria Urrutia
Lisa Blumerman	Nick Birnbaum
Kevin Deardorff	Sarah Brady
Kathleen Styles	

I. Comparison of Population Change: External Consultation and Current Research

John Long presented to the committee the demographic analysis (DA) work that has occurred since the March recommendation and is currently underway to explain the difference between DA and the A.C.E. results. Three meetings were held at the end of March with outside experts to examine the components of population change. These consultations resulted in the following recommendations for research:

- Review historical birth and death data and estimates
- Focus on components related to international migration
  - < Unauthorized Migration
  - < Emigration
  - < Temporary Migration
- Evaluate foreign-born data benchmarks
  - < Reweighted March 2000 CPS
  - < Census 2000 Supplementary Survey
  - < Preliminary data from Census 2000
- Validate components used in DA estimates 1990 to 2000.

John Long then updated the committee on the progress of the research to be presented to the ESCAP. The data for the 2000 foreign born population and the validation of the 1990 base are expected to be complete by mid-July. The data for the revision of the 1990 to 2000 components and the revised demographic analysis estimates are expected to be complete in August.

II. Next Meeting

The next meeting is scheduled for July 2, 2001 at 10:30. The agenda is to discuss E-Sample Erroneous Enumerations.

ESCAP MEETING NO. 52 - 07/02/01

AGENDA

Kathleen P Porter
07/02/2001 08:58 AM

To: Angela Frazier/DMD/HQ/BOC@BOC, Barbara E Hotchkiss/DSD/HQ/BOC@BOC,
Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Vanessa M Leuthold/DMD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC
cc: Roxanne Feldpausch/DSSD/HQ/BOC@BOC
Subject: Agenda for today's ESCAP meeting

The agenda for today's ESCAP meeting scheduled from 10:30-12 in Rm. 2412/3 is as follows:

E-Sample Erroneous Enumerations in A.C.E. - Roxanne Feldpausch

ESCAP MEETING NO. 52 - 07/02/01

HANDOUTS

Definition of E-Sample Universe: 1990 vs 2000.

	1990		2000	
	Number	Percent	Number	Percent
Census Count	248709873	100.0%	281421906	100.0%
Institutional, Military, etc (PES excluded)	4042361	1.6%	4694721	
Other GQ (ACE excluded)	2655383		3083912	
Total Non-Housing Unit (GQ) Population	6697744		7778633	2.8%
Housing Unit Pop	242012129	97.3%	273643273	97.2%
Imputed Pop Count	53655	0.0%	1172144	0.4%
Other whole household imputations	1547101	0.6%	2269030	0.8%
Other whole person substitions	300652	0.1%	2333092	0.8%
Data Define HU Population	240110721	96.5%	267869007	95.2%
"Late Census Adds" (DD)	240111		2239827	
Remote Alaska (DD)	43477		48503	
E-Sample Universe	**242482517**	**97.5%**	**265580677**	**94.4%**

Note: The numbers in this chart represent the approximate magnitude of the population of interest. However, not all numbers, especially those for 1990, have been verified and documented.

DRAFT: June 22, 2001

ESCAP MEETING NO. 52 - 07/02/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 52

July 2, 2001

Prepared by: Nick Birnbaum

The fifty-second meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on July 2, 2001 at 10:30 am. The agenda for the meeting was to discuss E-sample erroneous enumerations.

Committee Attendees:

Ruth Ann Killion
Cynthia Clark
John Thompson
Jay Waite
Howard Hogan
Nancy Potok
Nancy Gordon
Teresa Angueira
Bob Fay
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Raj Singh
Nick Birnbaum
Bill Bell
Danny Childers
David Whitford
Rita Petroni
Donna Kostanich
Tommy Wright
Roxanne Feldpausch

I. Changes in Definition of E-Sample Universe and Procedural Changes: 1990 vs. 2000.

John Thompson announced the topic for the meeting, which was to examine the results of the A.C.E. measurement of erroneous enumerations. The purpose of this presentation is to describe the components of Census 2000 erroneous enumerations, compare these components to the 1990 census, and to identify potential issues that require additional review and analysis. The next step in the review of erroneous enumerations will be the examination of the results of *The Analysis of Measurement Error Study*, which will analyze how accurately the A.C.E. measured erroneous enumerations. John Thompson then turned the meeting over to Howard Hogan who presented the Committee with data defining the E-sample universes for both 2000 and 1990 (see attached handout). These data were presented so that when the data on E-sample erroneous enumerations as measured by the A.C.E. were provided, Committee members would understand that because of changes in the universes, as well as procedural changes, the data on erroneous enumeration rates for the two censuses are not entirely comparable. How erroneous enumerations are categorized and some of the changes from 1990 to 2000 are discussed below.

- To be included in the E-sample, a person has to be data-defined. For Census 2000, for a person to be considered data-defined, we had to have at least two characteristics for that person, where name would count as a characteristic. For the 1990 census, "data-defined" required two characteristics, but name did not count as a characteristic, since it was not captured. E-sample erroneous enumerations fall into five categories: duplicate, fictitious, geocoding error, insufficient information for matching and followup, and other residence. These categories are defined below. Additionally, for some people, there was not enough information in the A.C.E. person followup interview to determine either their match or residence status. These people are called unresolved and had their probability of correct enumeration imputed.

  Duplicate – The census counted the same person more than once within the search area.

  Fictitious – The E-sample nonmatch was determined to be fictitious in this cluster during the A.C.E. person followup interview.

  Geocoding error – The housing unit exists outside the search area.

  Insufficient information for matching and follow-up – To have sufficient information for matching and followup, an E-sample person had to have a complete name and at least two characteristics. Included in this category were people with a blank or invalid name, or for whom we had only one characteristic.

Other residence – The A.C.E. person follow-up interview determined that the E-sample person was not a resident on Census Day because the person should have been enumerated at the other residence (includes died before Census Day and born after Census Day).

- Non-institutional, non-military group quarters were in the E-sample universe in 1990, but were excluded from the universe in 2000.

- In Census 2000 (unlike in 1990), an operation was conducted to remove duplicates. As a result of the unduplication operation, one would expect there to be fewer erroneous enumerations that were duplicates.

- There was a change in the search area for duplicates. As a result, some cases that would have been considered duplicates in 1990 would be coded as erroneous enumerations due to incorrect residence in 2000.

- A somewhat different set of cases were classified unresolved and imputed in the A.C.E. than in the PES.

II. Results on Erroneous Enumeration Rates

Some of the findings included:

- 1.8 percent of the E-sample people had insufficient information for matching and followup. It cannot be determined whether these cases were enumerated in error; however for the purpose of the dual system estimator, they are treated in the same way as those determined to be enumerated in error. Insufficient information was noticeably higher for enumerator-filled returns, especially in the case of proxy respondents. Of the people with insufficient information, 30.3 percent of these were designated as such because of an invalid name (e.g., Mickey Mouse or Donald Duck). There was very little geocoding error (0.2 percent), due in part to the way in which the Targeted Extended Search was defined.

- Higher erroneous enumeration rates were present for those race/origin domains that have higher undercounts. Non-Hispanic Whites had a lower rate than all other categories except American Indians on reservations.

- Owners had lower overall erroneous enumeration rates than non-owners, as expected.

- As expected, areas with high mail return rates had lower erroneous enumeration rates than areas with low mail return rates (and non-mail return areas).

- The following table compares the percent distribution of types of erroneous enumerations in 1990 and 2000:

Comparison Between 1990 and 2000 of the Percent Erroneous Enumerations

	2000	1990
Duplicate	0.8	1.6
Fictitious	0.3	0.2
Geocoding Error	0.2	0.3
Other Residence	1.0	2.2
Insufficient Information	1.8	1.2
Unresolved	0.6	0.3
Total	**4.7**	**5.8**

Note: 1990 data are from Hogan (1993) and related to the PES universe.

- It was noted from the table that the percentage of other residence erroneous enumerations dropped from 2.2 percent in 1990 to 1 percent in 2000. This was of concern to the Committee, and further analysis must be carried out to explain this difference.

III. Follow-up Examination of Issues Relating to Erroneous Enumerations

John Thompson requested that Howard Hogan's staff prepare a brief summary of some of the differences in erroneous enumeration rates between 1990 and 2000, including data comparing the percent distribution of types of erroneous enumerations (including insufficient information cases) by race.

In concluding the meeting, John Thompson raised the following issues for the Committee's consideration:

- To what extent do the results examined demonstrate real differences in census quality or are the differences largely due to methodological and definitional changes?

- This analysis focused on the measurement of erroneous enumerations identified in the E-sample. In order to fully evaluate all the data, it will be necessary to examine the P-sample findings as well.

- The next important step regarding the analysis of erroneous enumerations will be to examine the results of *The Analysis of Measurement Error Study*.

IV. Next Meeting

The next meeting is scheduled for July 11, 2001. The agenda for that meeting is to discuss the evaluation of P-sample non-matches.

ESCAP MEETING NO. 53 - 07/11/01

AGENDA

Kathleen P Porter
07/10/2001 09:53 AM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Teresa Angueira/FLD/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC
cc:
Subject: Agenda for 7/11 ESCAP meeting

The agenda for the July 11 ESCAP meeting scheduled from 10:30-12:00 in Rm. 2412/3 is as follows:

P-Sample nonmatches - DSSD

ESCAP MEETING NO. 53 - 07/11/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 53

July 11, 2001

Prepared by: Sarah Brady

The fifty-third meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on July 11, 2001 at 10:30. The agenda for the meeting was to discuss P-sample Nonmatches in the A.C.E.

Committee Attendees:

Nancy Potok
John Thompson
Cynthia Clark
Nancy Gordon
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
Ruth Ann Killion
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines
Bill Bell
Tommy Wright
Raj Singh
Donna Kostanich
Danny Childers
Glen Wolfgang
Kathleen Styles
Rita Petroni
Fay Nash
Maria Urrutia
Sarah Brady

I. P-sample Nonmatches

Glenn Wolfgang presented an analysis of P-sample nonmatches. The undercount may be viewed as the number of P-sample nonmatches balanced by the number of census erroneous enumerations, late census adds, and whole person imputations. It is important for the Committee to consider the nonmatches in light of how they may affect the overall undercount.

Glenn first presented the percent of resolved nonmover nonmatches by nonmatch household type (partial household nonmatched, whole household nonmatched in a matched housing unit, and whole household nonmatched in a nonmatched housing unit) for 1990 and 2000. The results are found in the following table:

Nonmatch Household Types: Percent of Resolved Nonmovers Not Matched for 2000 and 1990

Nonmatch Household Type	2000 Rate		1990 Rate	
	% P-Sample	%Non-match	% P-Sample	% Non-match
Partial Household Nonmatched	2.2	30.0	1.8	30.4
Whole Household Nonmatched in a Matched Housing Unit	3.3	45.9	2.3	38.5
Whole Household Nonmatched in a Nonmatched Housing Unit	1.7	24.1	1.8	31.1
TOTAL	7.2	100	5.9	100

The Committee determined that the overall increase from 1990 to 2000 in percent of nonmatches within the P sample may be due to the increase of late adds and whole person imputations. The Committee also noted that improved housing unit coverage may also explain the increased share of matched housing units.

Glenn then presented the P-sample nonmatch rates by post-strata. The overall nonmatch rate for 2000 was 8.2 percent, as compared to 7.8 percent for 1990. The following table presents the nonmatch rates by age/sex categories:

Nonmatch rates (%) by Age and Sex

	2000		1990	
Age/Sex	**Rate (%)**	**s.e.**	**Rate(%)**	**s.e.**
0-17	8.8	0.2	8.5	0.3
18-29 Male	13.2	0.3	13.3	0.4
18-29 Female	11.1	0.2	11.6	0.3
30-49 Male	8.5	0.2	7.9	0.3
30-49 Female	6.9	0.1	6.2	0.2
50+ Male	6.2	0.2	4.8	0.2
50+Female	5.6	0.1	4.1	0.2
TOTAL	8.2		7.8	

Note: 2000 data were computed from nonmovers and outmovers;
Official nonmatch rates used inmover data under certain conditions and may be slightly higher.
No change in test results was expected.
1990 data were computed from nonmovers and inmovers by Jim Liu and Lynn Imel (Mover Evaluation)

For the post-stratification levels of race/Hispanic origin domain, age/sex, and tenure, the percent of E-sample erroneous enumerations, percent whole person imputations, percent late census adds, and percent of net undercount were also presented. The Committee noted that (1) the distribution of nonmatches in 2000 appeared to be more similar to 1990 than the distribution of erroneous enumerations in 2000 was to1990 and (2) nonmatches are impacted by whole person imputations and late census adds, both of which have increased significantly from 1990. Therefore, the Committee stated that it was important to examine the characteristics of the imputations and late adds to more fully understand their relationship to the 2000 coverage measurement.

II. Next Meeting

The next meeting is scheduled for July 16, 2001 at 10:30. The agenda for the next meeting is to discuss conditioning.

ESCAP MEETING NO. 54 - 07/16/01

AGENDA

There was no agenda developed or used for the July 16, 2001 meeting.

ESCAP MEETING NO. 54 - 07/16/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting #54

July 16, 2001

Prepared by: Nick Birnbaum

The fifty-fourth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on July 16, 2001 at 10:30 am. The agenda for the meeting was to examine: 1) the evidence of contamination of Census 2000 data collected in A.C.E. block clusters and 2) an analysis of non-matches and erroneous enumerations using logistic regression.

Committee Attendees:

Ruth Ann Killion
Cynthia Clark
John Thompson
Jay Waite
Bob Fay
Howard Hogan
John Long
Carol Van Horn
Teresa Angueira

Acting Director/Deputy Director:
William Barron

Other Attendees:

Donna Kostanich
Dan Childers
Rita Petroni
Kathleen Styles
Sarah Brady
Michael Beaghen
Bill Bell
Tommy Wright
Dave Whitford
Nick Birnbaum
Raj Singh
Anne Kearney
Fay Nash

I. Contamination of Census 2000 Data Collected in A.C.E. Block Clusters

DSSD staff provided preliminary data and analysis from the evaluation study on contamination bias. The dual system estimation methodology assumes independence between the census and A.C.E. Failure of this assumption is sometimes referred to as contamination. Contamination or causal dependence occurs when the event of an individual's capture or non-capture in the initial census or the A.C.E. affects the probability of his or her capture in the other system.

The Census Bureau conducted an analysis to determine if the A.C.E. data collection activities contaminated Census 2000 data in the A.C.E. sample blocks. This analysis aggregates census data in the A.C.E. blocks to the national, evaluation poststrata, or regional and type of enumeration area (TEA) level. We then compare these data to census data in non-A.C.E. blocks, similarly aggregated, to see if significant differences exist. Using standard statistical tests, PRED examined evidence of contamination for three fundamental indicators, and for demographic, geographic, and response-related indicators.

While the analyses did reveal some evidence of extremely weak contamination in a few evaluation poststrata, regions, and TEAs, globally, there was not evidence of systematic contamination bias. Consequently, as additional loss function analyses are conducted, incorporating Census 2000 estimates for error components, the total error model will not include a component for contamination bias. It should be noted that these results are consistent with research undertaken on the 1990 census and test censuses leading up to Census 2000 that mostly show that the Census Bureau has not experienced contamination between the census and the corresponding coverage measurement survey.

II. Analysis of Non-Matches and Erroneous Enumerations Using Logistic Regression

DSSD staff then provided a preliminary analysis of census omissions and erroneous enumerations using logistic regression. The purpose of the logistic regression analysis was to identify those variables or characteristics that were important to or had the greatest influence on the likelihood of a person being matched or an erroneous enumeration. This analysis was designed to supplement previous presentations to the ESCAP on the components of E-sample erroneous enumerations (July 2, 2001) and P-sample non-matches (July 11, 2001). Those previous presentations focused on univariate analysis. The logistic regression analysis, by fitting a multi-variate model, allows one to parse out the effects of individual variables as well as to examine interactive effects. The effects of independent variables often seem stronger in univariate analyses, because the correlations among the variables are not taken into account. This analysis confirmed the results from the earlier presentations. That is, variables like tenure and race/ethnic origin that were observed to be important indicators in the prior analyses demonstrated similar relationships here, but the effects are weaker than they appear in the univariate analyses, in this case, because of the correlation between tenure and race/ethnic

origin. Therefore, the previous conclusions regarding the characteristics of nonmatches and erroneous enumerations remained in effect.

In addition, results from the P-sample modeling are evidence of the efficacy of the A.C.E. poststratification. The poststratification scheme was designed such that the probability of capture in the census would be as homogeneous as possible within each poststratum. The poststratification variables used were: age, sex, race/ethnic domain, tenure, region, and metropolitan statistical area/type of enumeration area (mail return rate was also a poststratification variable, but was not used in the model.). All of these variables were demonstrated in the model to be strong predictors of those who will be missed by the census.

III. Next Meeting

The agenda for the next meeting, scheduled for July 26, 2001, is to examine the 2000 nativity data and Census 2000 imputations.

ESCAP MEETING NO. 55 - 07/26/01

AGENDA

Kathleen P Porter
07/19/2001 09:49 AM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC,
Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Vanessa M Leuthold/DMD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Teresa Angueira/FLD/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC

cc: Violeta Vazquez/DMD/HQ/BOC@BOC, David A Raglin/PRED/HQ/BOC@BOC,
Elizabeth A Krejsa/PRED/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Michael J Batutis Jr/POP/HQ/BOC@BOC, Rosalyn R Harrington/DMD/HQ/BOC@BOC

Subject: ESCAP Meetings for Next Week

Please note on your calendars the ESCAP meetings and agendas scheduled for next week (both will be 10:30-12 in Rm. 2412/3):

July 26 (originally scheduled for July 25)	Imputation - DMD DA: 2000 Foreign Born Data - POP
July 27	Measurement of CEs and EEs w/o variances - Krejsa/Raglin PRED

ESCAP MEETING NO. 55 - 7/26/01
HANDOUTS

August 10, 2001

CENSUS 2000 INFORMATIONAL MEMORANDUM NO. 110

MEMORANDUM FOR	Preston J. Waite Associate Director for Decennial Census
From:	Teresa Angueira (Signed) Chief, Decennial Management Division
Subject:	Initial Research on Count Imputation in Census 2000
Contact Person:	Fay F. Nash, Assistant Division Chief for Statistical Design/Special Census Programs, Decennial Management Division, Room 2008-2, (301) 457-8039

A total of 1,172,144 persons, or .42 percent of the total population, was added to the apportionment count in Census 2000 through count imputation. While this rate was in line with earlier censuses, it was higher than the rate of count imputation in the 1990 Census. Accordingly, an interdivisional team was established to investigate and document the reasons for this occurrence. The team found that a variety of reasons attributed to the higher rate. This memo documents their initial findings and reflects the information presented to the Executive Steering Committee on Accuracy and Coverage Evaluation (A.C.E.) Policy at its July 26, 2001 meeting. A more detailed memorandum that will also include information on whole person characteristics imputation will follow upon completion of the research.

Background

The Census Bureau used count imputation in Census 2000 as it has in several prior censuses to address the problem of missing, incomplete, and contradictory data. The Census Bureau used count imputation for three categories of cases in Census 2000:

- Household Size Imputation – The Census Bureau imputed a population count for a housing unit when Bureau records indicated that the housing unit was occupied, but had insufficient information as to the number of individuals residing in the unit.

- Occupancy Imputation – When Census Bureau records indicated that a housing

unit existed but did not provide sufficient information to definitively classify it as either occupied or vacant, the Bureau imputed occupancy status (occupied or vacant). For a unit imputed as occupied, household size was also imputed.

- Status Imputation – When the Census Bureau's records had insufficient information about whether an address represented a valid, non-duplicated housing unit, the Bureau imputed the status of the unit (occupied, vacant, or delete). For a unit imputed as occupied, household size was also imputed.

As is shown in Table 1, the number of housing units subject to each of these three categories of count imputation was roughly equal.

Table. 1 Census 2000 Housing Units That Were Imputed in the Count Imputation Process by Category

	Number of Housing Units
Total	**620,650[1]**
Status Imputation	235,071 (38%)
Occupancy Imputation	191,826 (31%)
Household Size Imputation	193,753 (31%)

The team's research reveals that the explanations as to why more housing units were handled by the imputation process in Census 2000 than in 1990 vary by category of count imputation.

1. Status imputation

Status imputation in Census 2000 contributed to 235,071 housing units imputed as occupied or vacant, resulting in the imputation of 415,892 persons. The vast majority of the imputed housing units (97%) were No Return cases, with the remaining 3% being Enumerator Return cases. The Enumerator Return cases required status imputation because the questionnaires contained inconsistent information as to whether the unit should be classified as occupied, vacant or delete.

The No Return cases were those cases which were included on the Decennial Master Address File (DMAF) at the end of the census, but for which no data record was associated. Research revealed that 176,832 units (75% of the entire status imputation category) were census adds, meaning that they were housing units that were not on the DMAF by the time questionnaires were mailed out or delivered, but rather were added either by enumerators during field operations or by respondents themselves.

[1]Excludes cases with an imputed "Delete" status and thus removed from the census.

The team believes there are two possible reasons why we have no data for these added cases - the Non Identification (Non-ID) process and the constraints on the data processing schedule. The Non-ID process is a method by which new addresses are added to the census address file. A unique identifier, the MAF Identification (ID), is preprinted on questionnaires mailed or delivered to addresses for mailback, and preprinted on questionnaires given to enumerators for addresses they are assigned to visit during the field followup operations. However, enumerators working in the field may find previously missed addresses that need to be added to the census - addresses that do not have MAF IDs pre-assigned. Additionally, respondents may have completed Be Counted forms or responded through telephone questionnaire assistance (TQA), both of which would result in respondent data without a MAF ID. These added cases are assigned a temporary processing ID. The Non-ID process then matches these addresses to the MAF and assigns a MAF ID. The temporary processing ID links the newly assigned MAF ID to the appropriate data capture record. For some cases, the temporary processing ID was corrupted on the transmittal file of adds that entered the Non-ID process, thus preventing us from linking back to the corresponding data capture record.

Constraints on the data processing schedule was also a source of missing data for census adds. Some MAF IDs were entered on the DMAF very late in the census process - after the headquarters processing activities had begun. This is particularly true for the census operations conducted later in the census schedule, such as for the Coverage Improvement Followup (CIFU) operation. These late additions were not included in the merge process, whereby data captured records were merged with corresponding MAF IDs, that occurred at the beginning of the headquarters processing. Consequently no data records were associated with these late adds/MAF IDs. We only discovered this problem after the census counts had been released.

In summary, 75% of the housing units imputed in the status imputation category were added during the census enumeration process. They reflect valid housing units that were added either by enumerators in the field or by respondents themselves. These cases required imputation because we could not associate the corresponding data records to their final census ID numbers. However, these cases were appropriately included in the census.

The remaining 50,674 No Return cases (comprising 22% of the total status imputation category), consisted of census records that were data captured, but which contained no data, i.e., blank census records.

2. Occupancy Imputation

Occupancy imputation in Census 2000 contributed to 191,826 housing units imputed as occupied or vacant, resulting in the imputation of 260,652 persons. Of the housing units that required occupancy imputation, 179,149 (93%) were Enumerator Returns, 12,175 (6%) were in the No Return category, with the remaining 502 (0%) comprising the Mail

Returns.

The Enumerator Return cases required imputation because of inconsistent data recorded on the questionnaires, which precluded a definitive classification of either occupied or vacant. For example, the interviewer summary items may indicate the unit is vacant, but person data is recorded on the questionnaire. In Census 2000 no clerical edit process was implemented to resolve such inconsistencies prior to data capture as was done in the 1990 census. Instead, interviewer inconsistencies were handled by assigning an occupancy status via the automated imputation process, leading to a more standardized process.

The No Return cases were determined to be census adds verified in the Field Verification operation to exist as separate housing units, but for which no data capture record could be associated or for which the only record data captured was blank. The Non-ID process and constraints on the data processing schedule contributed to this group of cases, as well as to those cases described under the status imputation category (see above for more detail.)

3. Household Size Imputation

Household size imputation contributed to 193,753 occupied housing units in Census 2000 with population counts imputed**,** resulting in the imputation of 495,600 persons for this category. Of the housing units that were imputed, 159,761 (83%) were from Enumerator Returns, 29,402 (15%) were from Specialized Returns (such as Individual Census Returns, Individual Census Questionnaires, Military Census Returns, and Shipboard Census Returns), and 4,590 (2%) were in the No Return category.

The Enumerator Return cases required imputation because although the census record clearly indicated the unit was occupied, there was insufficient information about the household size. As with the Enumerator Returns requiring occupancy imputation, the higher rate of cases imputed under this category than in the 1990 Census can be substantially explained by the fact that for Census 2000 we did not perform a clerical coverage edit prior to data capture as was done in the 1990 Census.[2] Inconsistent or missing data caused these cases to be included in the count imputation process.

The Specialized Returns are single-person data collection forms. When these are the only forms data captured for a MAF ID, the ID does not necessarily represent a single-

[2]The number of this variety of count imputations would have been higher but for the fact that during the mid stages of the Non-response Followup operation, the Census Bureau discovered that a fairly large percentage of questionnaires were being processed with no population count. In order to reduce the number of potential person imputations, we implemented a process to identify these cases and send them back to the field to retrieve the missing information.

person household. Consequently, household size is imputed for such housing units.

The No Return cases are those for which the occupied status has been verified through a field operation, but for which no information is available from a census data record, i.e., a blank census record.

Conclusion

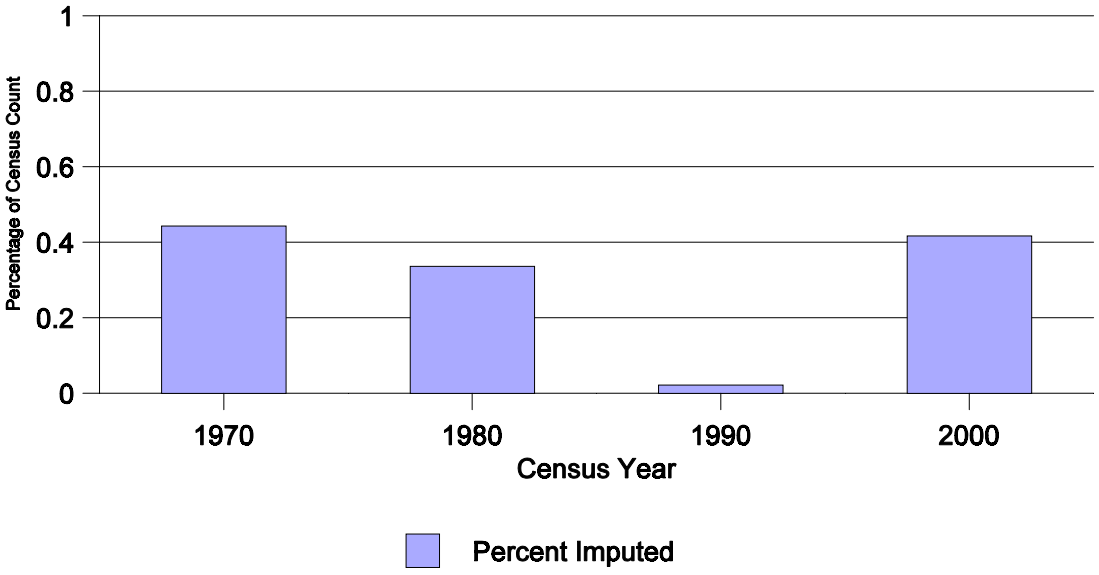
The team's research confirms that most of the count imputations performed in Census 2000 are attributable to housing units that have been determined to exist, but whose data were not included in the totals through a variety of reasons. These cases have been appropriately included in the census. If they had not been included in the count imputation process, these cases would represent individuals or housing units that should have been included in the Census, but who were left out because of incomplete or inconsistent data or the inability to locate appropriate data records due to processing system issues.

Attachment

cc: Team Members
Barbara Tinari (DMD)
Monique Sanders
Jane Ingold
James Treat (DSSD)
Nicholas Alberti
Richard Griffin
Gail Leithauser (FLD)
Mike Weiler
Dennis Stoudt (DSCMO)
Charles Kahn
David Galdi (GEO)
Lawrence Bates
Howard Hogan (DSSD)
John Clark
Brian Monaghan (FLD)
Robert Marx (GEO)
Michael Longini (DSCMO)

Attachment

Count Imputation Rates in the Decennial Census



ESCAP MEETING NO. 55 - 07/26/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 55

July 26, 2001

Prepared by: Sarah Brady

The fifty-fifth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on July 26, 2001 at 10:30. The agenda for the meeting was to discuss the consistency of provisional 2000 nativity data compared to benchmarks and to discuss imputations.

Committee Attendees:

John Thompson
Nancy Gordon
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
Ruth Ann Killion
John Long

Other Attendees:

Bill Bell	Nolan Malone
Tommy Wright	Rita Petroni
Raj Singh	Roxie Jones
Donna Kostanich	Kathleen Styles
Nick Alberti	Fay Nash
Gregg Robinson	Maria Urrutia
Signe Wetrogan	Sarah Brady
Kevin Deardorff	

I. 2000 Nativity Data

John Long began the presentation by explaining how it related to the examination of Demographic Analysis (DA) results from the March recommendation. In March, the Population Division developed alternative DA numbers. This alternative DA estimate doubled the number of illegal immigrants. The reweighted March 2000 Current Population Survey (CPS) results for the foreign born population were used as a benchmark to gauge the reasonableness of this assumption. This presentation looks at the provisional Census 2000 long form data and the results of the Census 2000 Supplementary Survey (C2SS) for foreign born to determine if these results are consistent with the reweighted March 2000 CPS and confirm the assumption of doubling the illegal immigration population. The next step will be to use long form and C2SS data to recalculate the legal and undocumented immigration components of DA. John Long then turned the presentation over to Kevin Deardorff.

Kevin Deardorff presented estimates of the foreign born population for the reweighted March 2000 CPS, provisional Census 2000 long form data (with and without group quarters), the original March 2000 CPS, and the C2SS. Both the provisional long form and the C2SS were not found to be significantly different from the reweighted March CPS for the foreign born population.

II. Imputations

Fay Nash presented data on Census 2000 imputations, which she was in the process of formally documenting in a memorandum. Subsequent to the meeting, Fay finalized the memorandum; it is attached and accurately reflects the meeting's discussion.

III. Next Meeting

The next meeting is scheduled for July 27, 2001 at 10:30. The agenda for the next meeting is to discuss measurement of erroneous enumerations.

ESCAP MEETING NO. 56 - 07/27/01

AGENDA

Kathleen P Porter
07/19/2001 09:49 AM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC,
Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Vanessa M Leuthold/DMD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Teresa Angueira/FLD/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC
cc: Violeta Vazquez/DMD/HQ/BOC@BOC, David A Raglin/PRED/HQ/BOC@BOC,
Elizabeth A Krejsa/PRED/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Michael J Batutis Jr/POP/HQ/BOC@BOC, Rosalyn R Harrington/DMD/HQ/BOC@BOC
Subject: ESCAP Meetings for Next Week

Please note on your calendars the ESCAP meetings and agendas scheduled for next week (both will be 10:30-12 in Rm. 2412/3):

July 26 (originally scheduled for July 25)	Imputation - DMD DA: 2000 Foreign Born Data - POP
July 27	Measurement of CEs and EEs w/o variances - Krejsa/Raglin PRED

ESCAP MEETING NO. 56 - 07/27/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 56

July 27, 2001

Prepared by: Sarah Brady

The fifty-sixth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on July 27, 2001 at 10:30. The agenda for the meeting was to discuss measurement of erroneous enumerations in the A.C.E.

Committee Attendees:

John Thompson
Nancy Gordon
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
Ruth Ann Killion
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Bill Bell	Kathleen Styles
Betsy Martin	Fay Nash
Tommy Wright	Maria Urrutia
Donna Kostanich	Sarah Brady
Dave Hubble	
Rita Petroni	
Dave Raglin	
Elizabeth Krejsa	

I. Measurement of Erroneous Enumerations

In the ESCAP meeting #54 on July 16, 2001 and the ESCAP meeting #52 on July 2, 2001, the Committee discussed erroneous enumerations. We noted that additional analysis would be forthcoming from the *Analysis of Measurement Error Study*. Today's meeting was a presentation of the additional analysis on erroneous enumerations. Dave Raglin presented data on the measurement of erroneous enumerations in the A.C.E. The data presented are from the *Analysis of Measurement Error Study*. In January and February of 2001, E- and P-sample cases in a sample of 1/5th of the A.C.E. clusters were sent out to the field for the Evaluation Followup Interview (EFU). The EFU includes a followup interview that gathered data in an attempt to resolve residency and matching issues, similar to the production person followup interview. The EFU used experienced interviewers. It also asks more probing questions than the PFU about a person's move-in and move-out dates and specific residence situations, such as college students, vacation homes, and various group quarters situations. Clerical matchers then took the results from the EFU along with previous information collected during production from the person matching and followup operations and assign match codes.

The most notable result from the *Analysis of Measurement Error Study* was that the EFU classified as erroneous enumerations 2,827,414 (weighted) people who had been classified as correct enumerations in production. There were 908,385 (weighted) people classified by production as erroneous enumerations that were classified as correct enumerations by the evaluation. Thus, the change from correct enumerations to erroneous enumerations is not balanced with the changes from erroneous enumerations to correct enumerations. Since the change is not balanced, it indicates that production potentially under classified people as erroneous enumerations, which would have a significant impact on the DSEs. That is, the production A.C.E. undercounts are too high.

Dave then presented demographic characteristics for the cases that moved from correct enumerations to erroneous enumerations.

The Committee discussed several concerns about this study. One concern was the effect of the dependent coding. EFU coders could reject the EFU interview and accept the PFU information. Another concern was that while the EFU asked more probing questions on some items, it was less detailed on other items. Other issues, including those related to the degree of review and rules followed in determining which (PFU or EFU) to accept as correct, were also discussed.

Since the results of the evaluation have a potentially large impact on the DSEs, the Committee decided further analysis into these cases is needed. The Planning, Research, and Evaluation Division (PRED) and the Decennial Statistical Studies Division (DSSD) have developed a proposal to have the even more highly trained matching analysts at NPC directly review a

sample of the EFU and production cases. Matching analysts are matchers at NPC with many years of training in matching, some with over 20 years of experience. They also supervise and perform quality assurance for all the A.C.E. matching operations. This work will start the beginning of August and results will be presented to ESCAP at a future meeting.

II. Next Meeting

The next meeting is scheduled for August 1, 2001 at 10:30. The agenda for the next meeting is to discuss the measurement of correct enumerations and to discuss the results of the evaluation of the 1990 Demographic Analysis.

ESCAP MEETING NO. 57 - 8/01/01

AGENDA

Kathleen P Porter
07/26/2001 10:01 AM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Teresa Angueira/FLD/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC
cc: Mary R Kennedy/POP/HQ/BOC@BOC, J Gregory Robinson/POP/HQ/BOC@BOC, Michael J Batutis Jr/POP/HQ/BOC@BOC, Violeta Vazquez/DMD/HQ/BOC@BOC, Signe I Wetrogan/POP/HQ/BOC@BOC
Subject: August 1 ESCAP meeting

There will be one ESCAP meeting next week. The agenda for the August 1 ESCAP meeting scheduled for 10:30-12 in Rm. 2412/3:

Evaluation of 1990 DA Results - POP

ESCAP MEETING NO. 57 - 08/01/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting #57

August 1, 2001

Prepared by: Nick Birnbaum

The fifty-seventh meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on August 1, 2001 at 10:30 am. The agenda for the meeting was to examine: 1) the number of undetected census discrepant persons in the A.C.E. production matching operation and 2) research evaluating the 1990 demographic analysis (DA) estimates.

Committee Attendees:

Ruth Ann Killion
Cynthia Clark
John Thompson
Jay Waite
Bob Fay
Howard Hogan
John Long
Carol Van Horn
Teresa Angueira
Nancy Potok
Nancy Gordon

Other Attendees:

Marvin Raines
Bill Bell
Kathleen Styles
Sarah Brady
David Raglin
Elizabeth Krejsa
Donna Kostanich
Nick Birnbaum
Raj Singh
Gregg Robinson
Arjun Adlakha

I. A.C.E. Erroneous Enumeration Errors: Underestimate of Census Discrepant Persons

In this presentation, PRED staff provided additional preliminary data and analysis from *The Matching Error Study* and the Evaluation Followup (EFU) interview to examine the extent to which discrepant people went undetected in the production matching operation. Discrepant results are those errors other than "honest" mistakes by interviewers or respondents and include falsification.

A person was classified as discrepant during the production matching operation if three knowledgeable respondents in the Person Followup (PFU) interview (building manager, neighbor, etc.) indicated not knowing him or her. To detect those discrepant persons that went undetected in production matching, "total error" match codes were used. These codes are the best match codes resulting from two additional operations – *The Matching Error Study* rematch and the EFU interview. The matching error study is a rematch of the production data for a sample of the A.C.E. clusters, while the EFU is a personal visit reinterview that was conducted in January and February 2001 in the evaluation clusters. The EFU is similar in purpose to the PFU – it gathers information to resolve conflicts between the A.C.E. and the initial census and to determine residence status. The data from the evaluation clusters are weighted to the national level.

The results indicated that any potential misidentification of discrepant persons in production matching operations had a minimal effect on the dual system estimates.

II. Evaluation of 1990 DA Estimates

Gregg Robinson began his presentation with a brief summary regarding the DA research program. The demographic analysis research program is reexamining the historic levels of the components of population change to address the scenarios dealing with the possibility that the 1990 demographic analysis estimates understated the Nation's population and that demographic analysis did not capture the full growth between 1990 and 2000. A major area of research involves re-examining the estimates of international migration, including unauthorized migration, legal immigration, emigration, temporary migration, and migration from Puerto Rico. Utilizing sample data from Census 2000 and other information sources, we can assess the accuracy of our current assumptions regarding these components. Secondly, we will examine the assumptions underlying other demographic analysis components, namely the birth, death, and Medicare data. The results of these research efforts will lead to a systematic re-calibration of the historical components of change, for years before 1990 and the1990 to 2000 period. When recompiled, the revised components will produce new demographic analysis estimates of population and coverage in 1990 and 2000.

Gregg then discussed the validation of the 1990 DA estimates. This process involves reconstructing and recalculating the components to ensure that no computational errors

had been made; it does not involve reconsideration of the assumptions underlying the estimates. Secondly, he reviewed the assumptions underlying the data for two of the components. All of the component data will be reviewed for consistency and completeness, and revisions will be made to the components if appropriate, and the 1990 DA estimates would then be revised accordingly. For today's presentation, Gregg discussed the reviews and recommendations regarding revisions to the historical births and deaths components of the 1990 DA estimates, based on consideration of the underlying assumptions relating to those estimates.

Validation of 1990 DA Estimates - The 1990 DA estimates were reconstructed and recalculated, using the full set of demographic components (1935 to 1990). The current reconstructed DA estimate for 1990 represents a 0.02% decrease from the 1990 DA estimate used as the basis for the March 1, 2001 ESCAP recommendation.

Consideration of Revisions to Historical Births - It was recommended that the assumption regarding birth under-registration (that is, no change from the levels of 1964-68, when the last test of birth completeness was conducted) be changed to reflect completeness of registration increasing linearly from 99.2% in 1966 to 100% by 1985 (and continuing to 2000). This revision to the under-registration assumption lowers the number of adjusted births from 1966 to 1990 by approximately 414,000. The rationale for this recommendation is based on, among others, the following factors:

- Senior staff at the National Center for Health Statistics believe that birth registration has improved over the last several decades and are unaware of any state having incomplete registration.

- All births in hospitals are issued a birth certificate and electronically recorded. Only 35,000 births occurred out of hospitals in 1999, and two-thirds of these were in some type of birthing facility.

Consideration of Revisions to Deaths to Population Under Age 65 in 1990 - After reviewing these data and determining that no change is warranted regarding the assumption of under-registration of infant deaths, Population Division recommended that there be no revisions to this component of the 1990 DA estimate.

The Committee also discussed whether deaths to recent unauthorized immigrants might cause the number of deaths to the authorized population to be somewhat overstated.

Data and analysis regarding proposed revisions of other components of the 1990 DA population estimate will be presented at upcoming meetings. In particular, there is considerable work that is being done on the immigration and emigration components.

III. Next Meeting

The agenda for the next meeting, scheduled for August 13, 2001, is to discuss how the A.C.E. estimates are affected by census whole person imputations and whether they can explain any of the differences between those estimates and the DA estimates.

ESCAP MEETING NO. 58 - 08/13/01

AGENDA

Kathleen P Porter
08/09/2001 04:34 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Vanessa M Leuthold/DMD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC
cc: Michael J Batutis Jr/POP/HQ/BOC@BOC
Subject: ESCAP Meetings for next week

ESCAP meetings for next week are as follows (all in Rm. 2412/3):

8/13 10:30-12:00 Imputations - POP and DSSD

8/16 10:30-12:00 Alt Models of Missing Data - PRED
Unresolved Cases (including Conflicting HHs) - DSSD

8/17 10:30-12:00 Evaluation of the 1990 Demographic Analysis Results (Part 2) - POP

ESCAP MEETING NO. 58 - 08/13/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 58

August 13, 2001

Prepared by: Theresa Leslie

The fifty-eighth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on August 13, 2001 at 10:30. As previously discussed at ESCAP, the level of whole person imputations has increased significantly over the previous census (although it was consistent with prior censuses.) The purpose of this meeting was 1) to understand the demographic characteristics of the whole person imputations as compared to the population not imputed, and 2) to understand the impact of whole person imputations on the measure of the undercount.

Committee Attendees:

Cynthia Z.F. Clark
Nancy Potok
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
John Long

Other Attendees:

Tommy Wright	Rita Petroni
Bill Bell	Florence Abramson
Donna Kostanich	Fay Nash
Raj Singh	Theresa Leslie
Dawn Haines	
Signe Wetrogan	
Greg Spencer	
Gregg Robinson	

The Executive Steering Committee on Accuracy and Coverage Evaluation Policy (ESCAP) discussed imputations in the census on July 11 and 26. As discussed at those meetings, the level of whole person imputations for Census 2000 increased significantly over the 1990 census but was consistent with prior censuses. ESCAP decided it was important to:

- understand the demographic characteristics of the whole person imputations as compared to the population not imputed and

- understand the impact of whole person imputations on the measure of the undercount.

The purpose of today's meeting was to provide data on these issues.

I. How do the demographic characteristics of whole person imputations compare to the characteristics of the population not imputed?

Signe Wetrogan, POP, presented tables showing the characteristics of whole person imputations. The age, race, and sex characteristics of the population requiring some form of imputation was similar to the data-defined population with the exception of the age category under age 18.

Upon closer examination, the relatively higher percent of the population under age 18 in the imputed population is due to the high proportions of younger people in the within household imputation universe. A large proportion of the within household imputation universe reflects the large households (seven or more members) that were not accommodated by the 6-person mail-return questionnaire. Because, most often, we are assigning age to the sixth and seventh persons within these large households, we would expect to assign a high proportion to the under 18 age group.

II. What is the impact of whole person imputations on the population undercount?

Howard Hogan, DSSD, presented data looking at characteristics of imputations for selected A.C.E. post-stratum groups. Differences between the post-stratum groups would translate into different impacts on the resulting estimates of undercounts or overcounts. From the data, DSSD concludes that the demographic characteristics of the whole person imputations as compared to the population not imputed is not an issue of concern.

Conclusion of the meeting: The whole person imputations did not negatively impact the measure of the undercount. The demographic characteristics of the whole person imputations as compared to the population not imputed did not differ beyond expectations.

III. Next Meeting

The next meeting is scheduled for August 16, 2001 at 10:30. The agenda for the next meeting is to discuss alternative models of missing data and unresolved cases in the A.C.E.

ESCAP MEETING NO. 59 - 08/16/01

AGENDA

Kathleen P Porter
08/09/2001 04:34 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Vanessa M Leuthold/DMD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC
cc: Michael J Batutis Jr/POP/HQ/BOC@BOC
Subject: ESCAP Meetings for next week

ESCAP meetings for next week are as follows (all in Rm. 2412/3):

8/13 10:30-12:00 Imputations - POP and DSSD

8/16 10:30-12:00 Alt Models of Missing Data - PRED
Unresolved Cases (including Conflicting HHs) - DSSD

8/17 10:30-12:00 Evaluation of the 1990 Demographic Analysis Results (Part 2) - POP

ESCAP MEETING NO. 59 - 08/16/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 59

August 16, 2001

Prepared by: Nick Birnbaum

The fifty-ninth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on August 16, 2001 at 10:30 am. The agenda for the meeting was to discuss: 1) unresolved cases in A.C.E. person matching and 2) alternative models for handling missing data from the A.C.E. interviewing.

Committee Attendees:

John Thompson
Jay Waite
Carol Van Horn
Howard Hogan
Nancy Potok
Nancy Gordon
Teresa Angueira
Bob Fay
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Raj Singh
Nick Birnbaum
Bill Bell
Carolee Bush
Sarah Brady
Maria Urrutia
Tom Belin (via telephone)
Rita Petroni
Donna Kostanich
Danny Childers
Don Keathley
Jim Liu
Anne Kearney
Kathleen Styles

This meeting was scheduled to discuss the potential effects of missing data on the accuracy of the A.C.E. The ESCAP first received a presentation on the levels and sources of unresolved cases. With this background, the ESCAP then received a presentation on alternative models that could be used to address missing data.

I. Analysis of Unresolved Codes in A.C.E. Person Matching

DSSD staff presented data on the distribution of unresolved status cases for the P-sample and the E-sample and compared these data to the 1990 PES. Persons were coded as "unresolved" if: 1) the A.C.E. person interview did not collect sufficient information for matching and followup, that is, the P-sample person did not have a complete name and at least two characteristics, or 2) the A.C.E. followup interview did not collect sufficient information to resolve the match status, residence status, or the enumeration status. Among the findings reported were the following:

- In the A.C.E., 2.2% of P-sample persons had unresolved residence status (including 1.2% with unresolved match status), and 2.6% of E-sample persons had unresolved enumeration status. For the 1990 PES, the comparable figures were: 1.8% of P-sample persons had unresolved match status, and 1.3% of E-sample persons had unresolved enumeration status. (Note: In the draft report distributed to Committee members, this last figure was erroneously reported as 2.3%.)

- More than half of the P-sample unresolved status cases were those coded as insufficient information for matching and followup (1.2% of the P-sample, compared to 0.4% in the 1990 PES).

- There is a difference in the coding procedures used for the A.C.E. versus the PES that affects the unresolved rates for the E-sample. People who did not live at the sample address but had an incomplete Census Day address as indicated on the followup interview form, were coded as erroneous enumerations in the 1990 PES but as unresolved cases in the A.C.E. In 2000, these unresolved enumeration status cases constituted 0.4% of the E-sample and were imputed as erroneous enumerations at high probabilities.

- Outmover and proxy interview cases accounted for slightly more than one half of the P-sample persons who had unresolved residence status.

- Among follow-up cases, 15.6% of P-sample persons and 14.6% of E-sample persons had unresolved statuses. Many of these unresolved status cases were due to proxy interviews.

II. Alternative Missing Data Procedures

Missing data in the A.C.E. result from non-interviews or item non-response. As in all surveys, missing data were accounted for through the use of missing data procedures. A non-interview adjustment accounted for non-interviewed households. Characteristic imputation was used for the following post-stratification variables: race, ethnicity, tenure, sex, and age. Finally, probability imputation was used to impute unresolved residence, match, or enumeration status.

This study was designed to assess the effect on the dual system estimates (DSEs) of using alternative missing data procedures; that is, the range in the national level DSE based on different combinations of alternative procedures. The resulting range would provide some indication of how sensitive the DSEs are to changes in one or more of the missing data procedures.

Seven "reasonable" alternative missing data procedures were selected (see attachment). These included: four alternative probability imputation procedures, two non-interview adjustment procedures, and a late data procedure. Of the 128 possible combinations of using/not using these seven alternative procedures, 32 of them were randomly selected for study. For each of the 32 combinations of alternative procedures, a DSE was computed at the national level. The results showed a range in the DSEs that was larger than expected, given the range produced by a similar analysis done on the 1990 PES. Of the various alternative procedures, the logistic regression model appeared to be the single largest factor in altering the DSEs. The Committee determined that additional analyses would be required to explain these results, particularly the larger than anticipated range in the DSEs for the 32 combinations examined.

III. Next Meeting

The next meeting is scheduled for August 22, 2001. The agenda for that meeting is to discuss the effect of excluding late census adds from the A.C.E person matching.

ESCAP MEETING NO. 59 - 08/16/01

HANDOUTS

Table 1. Alternative Missing Data Procedures

Alternative Procedure	Description	Motivation for using the Alternative Procedure
Late Data	Assign non-interviewed household weights to late-arriving household interviews only; use the same late-arriving interview information in imputing for probabilities.	Late-arriving interview data may more accurately reflect non-interviews and persons with unresolved match, resident, and enumeration status
Logistic Regression	Assign resident, match, and enumeration probabilities to unresolved cases using a logistic regression model.	Logistic regression models are accepted methods for estimating probabilites. This is what was done in 1990.
Alternative Non-interivew Adjustment (NIA) Cell Definitions	Use different NIA cells[1]. These alternative cells were defined on type of basic address, race/ethnicity/tenure, census division, state within division, type of enumeration area, and household size.	Household characteristics may be more homogeneous within alternative NIA cells.
Nearest-Neighbour NIA Method	Add each non-interviewed household's (donor) weight to the nearest (in a specified sort) interviewed household's (donee) weight. Each donee would receive no more than one donor's weight.	More homogeneity may result between donor and donee household characteristics when compared to spreading weights over many interviewed households.
Non-ignorable Missingness for Enumeration Probability	Lower the imputed enumeration, match, and resident probabilities for the corresponding unresolved cases. See Belin (2001) for a discussion on the research and procedure for lowering the probabilities.	Enumeration, match, and resident rates using resolved-only cases may overstate the corresponding rates for unresolved cases.
Non-ignorable Missingness for Match Probability		
Non-ignorable Missingness for Resident Probability		

[1] Production NIA cells were defined on block cluster, type of basic address category, re-coded A.C.E. sample stratum, and state (Cantwell 2001).

ESCAP MEETING NO.60 - 08/22/01

AGENDA

Kathleen P Porter
08/16/2001 04:33 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC,
Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Vanessa M Leuthold/DMD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC
cc:
Subject: ESCAP Meetings for next week

Please note on your calendars the ESCAP Meetings for next week:

August 22 Late Census Adds - Raglin

August 23 CANCELLED

ESCAP MEETING NO. 60 - 08/22/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 60

August 22, 2001

Prepared by: Sarah Brady

The sixtieth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on August 22, 2001 at 10:30. The agenda for the meeting was to discuss late census adds.

Committee Attendees:

John Thompson
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Ruth Ann Killion

Deputy Director/Acting Director:
William Barron

Other Attendees:

Bill Bell	Art Cresce
Marvin Raines	Gregg Robinson
Tommy Wright	Signe Wetrogan
Dave Hubble	Kathleen Styles
Rita Petroni	Fay Nash
Dave Raglin	Maria Urrutia
Raj Singh	Nick Birnbaum
David Whitford	Sarah Brady
Chester Bowie	

I. Effect of excluding late census adds from the A.C.E.

During the Census 2000 processing, some people in the census were not included in the A.C.E. person process. These were people that were initially deleted from the census but later reinstated into the census after the A.C.E. person process had begun. They were initially deleted because their housing units were suspected of being duplicates of other census housing units. In March of 2001, Howard Hogan prepared a memorandum documenting the effect of excluding late census adds from A.C.E. The memo stated that if the reinstated people were a small percentage of the correct enumerations in the census or their A.C.E. coverage rate was similar to the A.C.E. coverage rate for census people included in A.C.E., then there is a minimal effect on the Dual System Estimates (DSEs). To validate this assumption, we looked at the proportion of census correct enumerations that match A.C.E. people for the reinstates compared to other census people and the number of correctly enumerated reinstated people relative to the number of other census people.

Dave Raglin described the additional research and presented the results. The research involved clerically matching the reinstated people collected in the A.C.E. and census in the evaluation clusters. The evaluation clusters are a sample of 1/5th of the A.C.E. clusters. When we clerically matched the reinstates, the matchers followed similar procedures as production. The reinstated people were determined to be one of three enumeration statuses:

- Matched to an A.C.E. person living in the cluster on Census Day – correct enumeration
- Erroneously enumerated (mostly duplicates of census people)
- Not found in the A.C.E. or census

The results of the matching are presented in the following table:

Enumeration Status	Estimate	Percent
Matched to A.C.E. - correct enumeration	558,448	25.4%
Erroneously enumerated	1,153,418	52.5%
Not found in A.C.E. or census - undetermined enumeration status	486,626	22.1%
Total	2,198,492	100.0%

Subsequent to the meeting, typos were discovered in the table for the percent not found in A.C.E. or census and the estimate for the total. The above table was corrected for those typos.

For the purpose of this analysis, the erroneous enumerations among the late census adds

do not affect the proportion of census correct enumerations that match A.C.E. people. However, the unresolved cases will affect this ratio because a portion of them are correct enumerations. During the production matching, people who were unresolved were included in the Person Followup operation. It was not possible to follow these people up during this evaluation. Therefore, we assumed several different correct enumeration rates in a sensitivity analysis to provide a range of the effect on the A.C.E. coverage rate for the people with undetermined enumeration status and for the effect on the overall A.C.E. coverage rate. For the sensitivity analysis, we initially looked at the effect on the coverage rates if 50, 70, or 90 percent of the undetermined were correct enumerations. The Committee noted that using the assumption of 90 percent had the largest impact on the A.C.E. coverage rate, resulting in a decrease of 0.12 percent from production. A 0.12 percent decrease in the A.C.E. coverage rate would result in an increase of about 0.12 percent in the estimate of the undercount.

The Committee noted that 90 percent was very unlikely and discussed using a 33 percent rate in the sensitivity analysis, since 33 percent of the resolved people were correct enumerations. Subsequent to the meeting, a 33 percent correct enumeration rate was analyzed. If 33 percent of the unresolved cases were correct enumerations, then the A.C.E. coverage rate with reinstated cases would differ from the A.C.E. coverage rate without reinstated cases by 0.02 percentage points.

The Committee concluded that the overall effect of excluding the reinstated census people from the A.C.E. on the DSEs was minimal based on the change in the A.C.E. coverage rate ranging from 0.02 to 0.12 percentage points. Note– Subsequent discussions about the range of the change in the A.C.E. coverage rate due to including reinstated people led to a more plausible range of 0.034 to 0.082.

II. Characteristics of late census adds

Signe Wetrogan presented the characteristics of the late census adds. The Committee noted that the late census adds were not unusual or unexpected. The next step in examining the late census adds to is examine the effect they have on the calculation of the undercount as measured by DA.

III. Next meeting

The next meeting is scheduled for August 27, 2001 at 10:30. The agenda is to discuss the matching error rates for the A.C.E.

ESCAP MEETING NO.61 - 08/27/01

AGENDA

Kathleen P Porter
08/23/2001 01:41 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Vanessa M Leuthold/DMD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC
cc: Susanne L Bean/PRED/HQ/BOC@BOC, Xijian Liu/DSSD/HQ/BOC@BOC, David A Raglin/PRED/HQ/BOC@BOC, Richard A Griffin/DSSD/HQ/BOC@BOC, Eric L Schindler/DSSD/HQ/BOC@BOC
Subject: ESCAP Meetings for next week

Please note on your calendars the ESCAP meetings for the week of August 27 (all held in Rm. 2412/3):

August 27 10:30-12:00 Matching Error Rates w/variances - Bean,PRED

August 28 10:30-12:00 Mover Analysis - Liu, DSSD
Outmovers - Raglin,PRED

August 29 10:30-12:00 Synthetic Error - Griffin/Schindler, DSSD

ESCAP MEETING NO. 61 - 08/27/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 61

August 27, 2001

Prepared by: Nick Birnbaum

The sixty-first meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on August 27, 2001 at 10:30 am. The agenda for the meeting was to discuss the evaluation of matching error in the A.C.E. as measured by the Matching Error Study (MES).

Committee Attendees:

Jay Waite
Carol Van Horn
Howard Hogan
Nancy Potok
Teresa Angueira
Bob Fay
Ruth Ann Killion

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines
Bill Bell
Rita Petroni
Donna Kostanich
Dave Raglin
Sarah Brady
Tommy Wright
Dave Hubble
Susanne Bean
Nick Birnbaum
Fay Nash
Kirsten West
Maria Urrutia

I. MES Evaluation of Matching Error in the A.C.E.

PRED staff presented data and analyses from the Matching Error Study (MES) relating to the level of matching error in the A.C.E. For the MES, an independent rematch was conducted in the A.C.E. evaluation clusters. The rematch was done using production procedures. If the rematch outcome differed from production, then analysts (the most highly trained matching personnel) were used to reconcile the differences.

The results indicated that the level of matching error was lower in 2000 than 1990 – evidence that the changes in procedures improved the quality of the matching operations. In the present analyses, a number of metrics were used to measure the differences in matching error between the 1990 PES and the 2000 A.C.E. For example, the study indicated that the gross difference and net difference rates for the P- and E-samples were lower in 2000 than in 1990. Another analysis looked at the relative bias in the number of P-sample matches and E-sample correct enumerations between the production and the MES figures. The role of the correct enumeration rate and the match rate in the calculation of the dual system estimate (DSE) was discussed. At the national level, the 2000 relative difference rate ((production-rematch)/rematch) for correct enumerations shows a reduction from the 1990 rate, and the 2000 relative difference rate for matches was similar to that from 1990. The relative difference rates for 1990 and 2000 were also examined for evaluation post-stratum groups. While these post-stratum groups were defined somewhat differently in the two coverage measurement surveys, for 2000, the results indicate smaller ranges in the relative difference rates among post-strata for both matches and correct enumerations. This finding is another indication of the reduction in matching error.

The study also examined the extent of clerical errors in identifying duplicates in the A.C.E. search area. The results of this analysis indicate that for both E- and P-sample cases, the numbers of false or missed duplicates were small and thus contributed minimally to the level of matching error in the A.C.E.

Finally, the impact of matching error on the DSEs was assessed by examining the difference (to measure error) between the production and MES ratios of the correct enumeration rate to the match rate for the sixteen evaluation post-strata and at the national level. Depending upon the statistical test used, it was noted that matching error significantly inflated the production DSEs for as many as five of the evaluation post-strata. Consequently, the national production DSE was approximately 400,000 higher than the MES DSE.

II. Next Meeting

The next meeting is scheduled for August 28, 2001. The agenda for that meeting is to discuss the analysis of movers and resident status in the A.C.E.

ESCAP MEETING NO. 62 - 08/28/01

AGENDA

Kathleen P Porter
08/23/2001 01:41 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC,
Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Vanessa M Leuthold/DMD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC
cc: Susanne L Bean/PRED/HQ/BOC@BOC, Xijian Liu/DSSD/HQ/BOC@BOC, David A
Raglin/PRED/HQ/BOC@BOC, Richard A Griffin/DSSD/HQ/BOC@BOC, Eric L Schindler/DSSD/HQ/BOC@BOC
Subject: ESCAP Meetings for next week

Please note on your calendars the ESCAP meetings for the week of August 27
(all held in Rm. 2412/3):

August 27 10:30-12:00 Matching Error Rates w/variances - Bean,PRED

August 28 10:30-12:00 Mover Analysis - Liu, DSSD
Outmovers - Raglin,PRED

August 29 10:30-12:00 Synthetic Error - Griffin/Schindler, DSSD

ESCAP MEETING NO. 62 - 08/28/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 62

August 28, 2001

Prepared by: Sarah Brady

The sixty-second meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on August 28, 2001 at 10:30. The agenda for the meeting was to discuss movers and changes in mover and residence status.

Committee Attendees:

Nancy Potok
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Ruth Ann Killion

Deputy Director/Acting Director:
William Barron

Other Attendees:

Bill Bell	Jim Liu
Marvin Raines	Dan Weinberg
Tommy Wright	Maria Urrutia
Dave Hubble	Nick Birnbaum
Rita Petroni	Sarah Brady
Dave Raglin	
Elizabeth Krejsa	
Donna Kostanich	
Dan Childers	

The purpose of this meeting was to discuss the analysis of movers from Census Day as measured by the A.C.E. There were two presentations on this topic: Jim Liu presented data on the level and type of movers and Dave Raglin presented results from the EFU about the accuracy of the handling of movers. Jim's presentation was background for Dave's presentation.

I. Analysis of movers as measured by the production A.C.E.

Jim Liu compared the percent of inmovers in 2000 to 1990. In 1990, the percent of inmovers was 7.8 percent; in 2000, it was 5.1 percent. Inmovers are people who lived at the housing unit on the A.C.E. interviewing day, but not on Census Day. Jim explained that the decrease from 1990 could be explained by the earlier interviewing dates for the 2000 coverage measurement survey. In 1990, the PES interviewing started on June 25. In 2000, A.C.E. interviewing began on April 24. The A.C.E. interviews were much closer to Census Day, which would reduce the number of movers.

Jim also presented mover match rates by mover status (nonmovers vs movers) and demographic characteristics. It was concluded that mover match rates in 2000 were generally higher than in 1990.

John Thompson asked Jim to quantify the amount of under/over-count due to movers. The Committee discussed methodologies to quantify the impact of movers. The Committee was unable to agree upon any readily available methodology. Subsequent discussions indicated that this calculation was unnecessary for the ESCAP deliberations.

II. Classification of mover and residence status as measured by EFU

Dave Raglin presented results from the Evaluation Followup (EFU) interview on the classification of mover and residence status. The EFU is a followup interview conducted in a sample of the A.C.E. clusters during the months of January and February 2001. The EFU asked questions to determine if a person was a resident of the housing unit on Census Day. The evaluation found there were more people changing from residents to nonresidents than vice versa. Many of these cases as part of the PFU/EFU review to evaluate the EFU results.

The results for changes in mover status are as follows:

- Nonmovers were consistent– production nonmovers remained nonmovers in EFU the majority of the time.
- Approximately 20 percent of the time, production outmovers were reclassified as nonmovers in EFU– No explanation for this reclassification has been developed; further research is needed to explain it.
- Over 30 percent of inmovers became nonmovers or outmovers– Approximately half of those people were either duplicates of existing A.C.E. people or matches to census people. In this case, we believe the people were really Census Day residents and either the respondent misidentified these people as having moved in since Census Day or the interviewer entered the data incorrectly on the laptop. An analysis of the

trace files is planned for after the October 15 recommendation date to verify this hypothesis.

The final report will address how the results of these studies affect the assessment of the accuracy of the A.C.E.

III. Next Meeting

The next meeting is scheduled for August 29, 2001 at 10:30. The agenda is to discuss synthetic error.

ESCAP MEETING NO. 63 - 08/29/01

AGENDA

Kathleen P Porter
08/23/2001 01:41 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, Vanessa M Leuthold/DMD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC
cc: Susanne L Bean/PRED/HQ/BOC@BOC, Xijian Liu/DSSD/HQ/BOC@BOC, David A Raglin/PRED/HQ/BOC@BOC, Richard A Griffin/DSSD/HQ/BOC@BOC, Eric L Schindler/DSSD/HQ/BOC@BOC
Subject: ESCAP Meetings for next week

Please note on your calendars the ESCAP meetings for the week of August 27 (all held in Rm. 2412/3):

August 27 10:30-12:00 Matching Error Rates w/variances - Bean,PRED

August 28 10:30-12:00 Mover Analysis - Liu, DSSD
Outmovers - Raglin,PRED

August 29 10:30-12:00 Synthetic Error - Griffin/Schindler, DSSD

ESCAP MEETING NO. 63 - 08/29/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting #63

August 29, 2001

Prepared by: Nick Birnbaum

The sixty-third meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on August 29, 2001 at 10:30 am. The agenda for the meeting was to discuss synthetic error in the A.C.E. estimates.

Committee Attendees:

Ruth Ann Killion
Jay Waite
Bob Fay
Howard Hogan
John Long
Carol Van Horn
Teresa Angueira
Nancy Potok
Nancy Gordon

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines
Bill Bell
Kathleen Styles
Sarah Brady
Rita Petroni
Fay Nash
Tommy Wright
Donna Kostanich
Nick Birnbaum
Maria Urrutia
Gregg Robinson
Eric Schindler
Rick Griffin

I. Sensitivity Analysis for the Assessment of the Synthetic Assumption

In report B-14*, Griffin and Malec assessed the level of error in synthetic estimates at the state and congressional district levels and the effect of this error on the loss function results. Those loss function results used one of eight sets of assumptions dealing with correlation bias and A.C.E. processing error (the DSE bias component) and one of two methods to synthetically distribute total error model targets to states and congressional districts. The purpose of the present analysis was to conduct sensitivity analysis on the synthetic error's effect on the loss function results by varying these eight assumptions and two methods.

The analysis examined loss function results for estimated state levels (counts), estimated state shares, and estimated congressional district shares. The results of this sensitivity analysis included the following:

- For estimated state levels, the analysis revealed no change in the loss function result (favoring the census counts or the A.C.E. estimates) for all of the ninety-six combinations of eight DSE bias models, two distribution methods, and six synthetic bias models (based on six artificial populations). That is, except for the combinations including DSE bias model #8 (no correlation bias and 100% of the 1990 census level of processing error in the A.C.E.), all other combinations favored the A.C.E. estimates.

- For estimated state shares, the sensitivity analysis showed a switch in the loss function results for eighteen of the 96 combinations, sixteen of which change the result from favoring the A.C.E. to favoring the census numbers. Still, for almost three-quarters of the combinations, the loss function results favor the A.C.E. estimates.

- For estimated congressional district shares, there is virtually no change, with the loss function results favoring the A.C.E. estimates in most cases; however, for the twelve combinations that include DSE bias model #8, the majority of these cases favor the census counts.

Subsequent review of the results generated a request for data describing the relative effect of synthetic error on each of the census and A.C.E. loss functions. Concerns were also expressed regarding whether the A.C.E. loss functions included the effects of sampling error.

II. Alternative Assessment of the Synthetic Assumption

An alternative assessment of synthetic error was presented by DSSD staff using direct state DSEs. However, several technical concerns were raised about the use of these direct state estimates to determine the level of synthetic error in the A.C.E. estimates. Consequently, this analysis will not be used by the Committee in its ongoing assessment of the data and analyses presented to it.

III. Next Meeting

The agenda for the next meeting, scheduled for September 5, 2001, is to discuss the results of the housing unit coverage study.

ESCAP MEETING NO. 64 - 09/05/01

AGENDA

Kathleen P Porter
08/30/2001 02:20 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC,
Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC
cc: Michael A Beaghen/DSSD/HQ/BOC@BOC, Diane F Barrett/DSSD/HQ/BOC@BOC
Subject: ESCAP Meeting for next week

Please note on your calendars the ESCAP meeting for the week of September 4:

September 5 10:30-12 Rm. 2412/3 HU Coverage Study - Barrett/Beaghen (DSSD)

ESCAP MEETING NO 64. 09/05/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 64

September 5, 2001

Prepared by: Sarah Brady

The sixty-fourth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 5, 2001 at 10:30. The agenda for the meeting was to discuss the housing unit coverage study.

Committee Attendees:

Nancy Potok
John Thompson
Cynthia Clark
Nancy Gordon
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
Ruth Ann Killion
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines	Michael Beaghen
Tommy Wright	Diane Barrett
Rita Petroni	Fay Nash
Joseph Burcham	Maria Urrutia
Raj Singh	Kathleen Styles
Donna Kostanich	Sarah Brady
David Whitford	
Dan Childers	

The purpose of this meeting was to examine the results of the housing unit coverage study. This study will provide information to the Committee about the quality of the A.C.E. data by comparing the results of housing unit coverage in 2000 to that in 1990.

I. Housing Unit Coverage Study

Diane Barrett presented the results of the housing unit coverage study. The study measures the housing unit coverage of Census 2000 using the A.C.E. housing unit data. The national percent undercount for housing units was 0.61. This is a decrease from the 0.96 percent undercount in 1990. The undercount rate for vacant units was higher than for occupied units, 3.37 and 0.33 percent respectively. The vacancy rate in Census 2000 has been a concern; other data series indicate that there were too few vacant units. The results from the housing unit coverage study are consistent with these other data series.

Diane Barrett then presented housing unit coverage rates by demographic characteristics such as tenure, race/Hispanic origin of person 1, type of structure, metropolitan statistical area/type of enumeration area. We did not perform tests for significant differences between types in these categories. Any comparisons mentioned below are based on the percent undercount only. The Committee noted the following interesting results:

- The housing unit coverage rates by type of structure for occupied units indicate that there was an overcount for small multi-units with 2 to 9 housing units. There was an undercount for single units. Some Committee members noted that this finding was surprising, since they were expecting an undercount for small multi-units with 2 to 9 housing units and an overcount for single units.

- When examining the coverage rates by metropolitan statistical area/type of enumeration area, the Committee noted that the size of the metropolitan statistical area didn't affect coverage.

- The most common reason for coding a housing unit as an erroneous enumeration was because the unit was not a housing unit (57.05%)-- the unit could be, for example, a group quarters or place of business. The second most common reason for coding a housing unit as an erroneous enumeration was because the unit was a duplicate (24.81%).

- The Committee discussed the housing unit coverage of the Black and Hispanic populations and ultimately concluded the data did not present evidence of undercounts given the size of the sampling error.

These data do not appear to be inconsistent with the 1990 Housing Unit Coverage Study or exhibit large inconsistencies with the person coverage results from A.C.E. No additional concerns were added to our recommendation process due to housing unit coverage.

II. Next Meeting

The next meeting is scheduled for September 12, 2001 at 10:30. The agenda is to discuss an evaluation of census person duplication.

ESCAP MEETING NO. 65 - 09/12/01

AGENDA

Kathleen P Porter
09/12/2001 09:29 AM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC,
Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, RJones17@doc.gov, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC
cc: Theresa F Leslie/DMD/HQ/BOC@BOC, Vincent T Mule Jr/DSSD/HQ/BOC@BOC
Subject: ESCAP MEETING TODAY AT 10:30

The ESCAP Meeting scheduled for yesterday on Census Person Dups will be held TODAY from 10:30-12:00 in Rm. 2412/3.

ESCAP MEETING NO. 65 - 09/12/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 65

September 12, 2001

Prepared by: Sarah Brady

The sixty-fifth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 12, 2001 at 10:30. The agenda for the meeting was to discuss an evaluation of census duplication.

Committee Attendees:

Nancy Potok
John Thompson
Nancy Gordon
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
John Long

Other Attendees:

Marvin Raines	Fay Nash
Bill Bell	Maria Urrutia
Tommy Wright	Theresa Leslie
Dave Hubble	Sarah Brady
Rita Petroni	
Raj Singh	
Donna Kostanich	
Debbie Fenstermaker	
Tom Mule	
Kirsten West	

I. Background– Census Person Duplication

One of the areas of concern for the Committee during the ESCAP process for the March 1 recommendation was that the 2000 A.C.E. by design did not measure duplication between certain components of the population enumerated in both group quarters and housing units. We were concerned that perhaps the estimate of erroneous enumerations in the 2000 A.C.E. was too low because the estimate of duplicate enumerations as measured by the A.C.E. was lower than the estimate from the 1990 Post-Enumeration Survey (PES). Our matching work identified duplicate enumerations that were outside of the scope of the A.C.E. This included duplicate enumerations identified outside of the geographic search area and enumerations in housing units and group quarters outside of the A.C.E. universe. Significant duplication of this type could explain some of the differences between demographic analysis and A.C.E.

An interdivisional team conducted a computer match between census cases in the A.C.E. sample clusters and the entire census, called the source and target files. The computer match was conducted in two stages. The first stage of matching was an exact match on first name, last name, month of birth, and day of birth. The housing units linked during the first stage of matching were then sent to the second stage. During the second stage of matching, all the person records from the linked source unit and from the linked target unit were statistically matched by comparing first name, middle initial, last name, month of birth, day of birth, and computed age. Then modeling was used to give a weight of duplication to each link.

This study addressed two major questions:

1. Why was the estimate of duplication in the 1990 Post-Enumeration Survey (PES) different than the 2000 A.C.E.?

2. What was the extent of duplicate enumerations that were 1) outside of the search area and 2) outside of the universe of A.C.E.?

Tom Mule presented the results of the evaluation and the answers to the questions stated above.

II. Why was the estimate of duplication in the 1990 Post-Enumeration Survey (PES) different than the 2000 A.C.E.?

The 1990 PES measured approximately 3.8 million duplicate enumerations or approximately 1.6 percent. However, the A.C.E. measured approximately 2 million or 0.8 percent duplicate enumerations. Also, there was a housing unit duplication operation in 2000 that was not a part of the 1990 census. This operation temporarily removed housing units from the census that were believed to be duplicates. These units were examined and some of them were determined to be duplicates and deleted and some were

determined to be distinct units and were reinstated into the census. It was known that some of the reinstated units contained duplicate people. However, a decision was made to reinstate duplicate persons because the housing units were unique. The reinstated units were excluded from the A.C.E. matching; the effect of the reinstated units was discussed at ESCAP meeting # 60. Excluding these units from the matching operation meant that the people in the units were not eligible to be searched for duplicates.

In addition, all group quarters were out-of-scope for the A.C.E., but the PES included non-institutional group quarters. To compare the duplication measured in 1990 and 2000, we would need to look at the amount of duplication between people in housing units and group quarters in 2000 and the duplication between people in reinstated housing units and the census. The census person evaluation estimated that there were 1,223,632 people duplicated in group quarters and reinstated units (149,904 in group quarters and 1,073,728 in reinstated units). Therefore, had the A.C.E. implemented the same methodology as the 1990 PES, it would have measured 3,238,307 duplicate people or approximately 1.2 percent. The large number of duplicates identified by this analysis as out-of-scope for the A.C.E., potentially explains a portion of the difference in duplicate enumerations between the 1990 PES and the 2000 A.C.E.

III. What was the extent of duplicate enumerations that were 1) outside of the search area and 2) outside of the universe of A.C.E.?

Tom Mule then presented results from the evaluation for the following universes: outside of the A.C.E. search area and outside of the A.C.E. universe. The evaluation found 2,089,456 duplicated people between census housing units outside of the surrounding blocks. This figure does not include duplicate links made to reinstated units. There were 660,219 people duplicated from census housing units to group quarters. To understand whether this level of duplication is significant and what its effect is on the A.C.E. estimates, we need to examine how the people identified in this evaluation as duplicated outside the A.C.E. cluster were coded in the A.C.E. Staff in the Decennial Statistical Studies Division (DSSD) are conducting this evaluation. The presentation for this evaluation is scheduled for September 26.

Tom Mule also presented tables and figures showing the duplication for two of the A.C.E. post-stratification variables: Race and Hispanic Origin domains and Age/Sex categories. He made the following points about duplication that was out of scope for the A.C.E. when looking at the Race/Hispanic Domain and the Age/Sex categories:

- For census housing units to census housing units, Non-Hispanic Blacks and Hispanics had higher percentages of duplication outside the surrounding blocks but still within the county than did those in the Non-Hispanic White or Some Other Race categories.

- There was greater duplication of Non-Hispanic Blacks than Hispanics between housing units and group quarters. Non-Hispanic Blacks had higher amounts of

duplication between: 1) housing units and correctional facilities, and 2) housing units and college dorms.

- There were higher estimates of duplication for the three age/sex categories under 30 than the four categories over 30. Duplication for the under 30 age categories was seen more often in the same county while duplication for the 50 plus age categories was seen more often in a different state.

- The 18-29 males and 18-29 females had higher amounts of duplication between housing units and group quarters than the other age/sex categories. The 18-29 female group was predominantly in college dorms while the 18-29 male group was duplicated in college dorms, correctional facilities and military group quarters.

- For census housing units to deletes, we saw no differences based on Race/Ethnicity domain or Age/Sex category.

These results were consistent with our expectation of where duplication would occur.

IV. Next Meeting

The next meeting is scheduled for September 13, 2001 at 10:30. The agenda is to discuss issues of balancing in the A.C.E.

ESCAP MEETING NO. 66 - 09/13/01

AGENDA

Kathleen P Porter
09/06/2001 02:38 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC,
Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC
cc: Theresa F Leslie/DMD/HQ/BOC@BOC, Vincent T Mule Jr/DSSD/HQ/BOC@BOC,
Tamara S Adams/DSSD/HQ/BOC@BOC
Subject: ESCAP Meetings for week of Sept. 10

The ESCAP Meetings for the week of September 10 are as follows (all are from 10:30-12:00 and in Rm. 2412/3):

September 10 CANCELLED

September 11 Census Person Duplicates - Mule/Leslie (DMD)

September 13 Balancing - Adams (DSSD)

ESCAP MEETING NO. 66 - 09/13/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting #66

September 13, 2001

Prepared by: Nick Birnbaum

The sixty-sixth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 13, 2001 at 10:30 am. The agenda for the meeting was to discuss the issue of balancing.

Committee Attendees:

John Thompson
Ruth Ann Killion
Cynthia Clark
Jay Waite
Bob Fay
Howard Hogan
John Long
Carol Van Horn
Teresa Angueira
Nancy Potok
Nancy Gordon

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines	Donna Kostanich
Bill Bell	Nick Birnbaum
Kathleen Styles	Maria Urrutia
Sarah Brady	Raj Singh
Rita Petroni	David Whitford
Fay Nash	Danny Childers
Tommy Wright	Tammy Adams

I. Balancing

Balancing error, specifically geographic balancing error, occurs when the effective search area for finding matches differs from the search area used to define correct enumerations. For example, assume that the searching for matches only looked within the sample block. Now if the estimation system counted as a correct enumeration a person counted outside the sample block, there would be two different definitions used for defining matches as for defining correct enumerations. The measurement process would not balance and would produce an overestimate of the net undercount. If the situation was reversed (extended looking for matches but narrow definition of correct enumerations), the result would be to underestimate the net undercount.

A symptom of balancing error would be far more matches in the surrounding block than cases considered correct because they were coded into the surrounding block. Such a situation was observed in the A.C.E. estimates, specifically there were approximately 3 million more matches into the surrounding block than were coded as correctly enumerated but in the surrounding block. In its March recommendation, the ESCAP was greatly troubled by this finding. However, because there are other explanations of this symptomatic lack of balance that do not result in a bias in the A.C.E. estimates, the Census Bureau conducted several studies. Specifically, in Targeted Extended Search (TES) 2, the following types of E-sample units were followed up: erroneously enumerated, adds coded as geocoding errors in TES-eligible clusters, and census units in list/enumerate clusters. These units could either be found in the surrounding blocks or beyond the surrounding block ring. If they were found in a surrounding block, this would increase the number of correct enumerations in surrounding blocks and decrease the lack of balance. If the people were already coded as correct enumerations, then the correct enumeration rate would not change. If these units were found beyond the surrounding ring, then they should have been coded as erroneous enumerations due to geocoding error, and the correct enumeration rate was higher than it should have been.

In TES3, there was an E-sample component and a P-sample component to the field follow-up work. The following cases of P-sample units were followed up: matches to the surrounding blocks, nonmatched housing units with nonmatched people, matched units where the census half of the unit was deleted in TES-ineligible clusters, the control sample, and other units that do not fall into the previous categories (e.g., conflicting households, whole household possible matches, and noninterviews). These units could either be found in the surrounding blocks or beyond the surrounding block ring. The units should not have been originally listed in the P-sample. However, the matches to the surrounding blocks will correct for an otherwise overstated nonmatch rate.

TES3 also followed up two types of E-sample housing units: correct or unresolved housing units and adds outside the cluster in TES-ineligible clusters. These units could either be found in the surrounding blocks or beyond the surrounding block ring. Similar to the analysis presented above, if these E-sample units were found in a surrounding

block, this would increase the number of correct enumerations in surrounding blocks and decrease the lack of balance. If the people were already coded as correct enumerations, then the correct enumeration rate would not change. Again paralleling the above analysis, if these units were found beyond the surrounding ring, then they should have been coded as erroneous enumerations due to geocoding error, and the correct enumeration rate was higher than it should have been.

Based on the results of TES2 and TES3, the estimate for the lack of balance between the P-sample person matches to surrounding blocks and the E-sample person correct enumerations in surrounding blocks is considerably smaller than the initial data indicated. Consequently, the concerns regarding the lack of balance were, for the most part, allayed. The level of geocoding error in the A.C.E. revealed in these results was determined to have a trivial effect on the dual system estimates. However, the TES2 and TES3 analysis did indicate that some A.C.E. errors are not entirely accounted for in the total error model.

II. Next Meeting

The agenda for the next meeting, scheduled for September 17, 2001, is to discuss the revised demographic analysis estimates.

ESCAP MEETING NO. 67 - 09/17/01

AGENDA

Kathleen P Porter
09/14/2001 02:09 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, RJones17@doc.gov, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, roxie.jones@mail.doc.gov, RJones17@doc.gov
cc: Donald H Keathley/PRED/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Mary Helen Mulry/SRD/HQ/BOC@BOC, Roxanne Feldpausch/DSSD/HQ/BOC@BOC
Subject: ESCAP Meetings for week of September 17

The following dates are scheduled ESCAP meetings for the week of September 17 (all are in Rm. 2412/3):

9/17 10:30-12:00 Remaining DA - POP

9/20 10:30-12:00 Missing Data - Keathley (PRED)
Correlation Bias - Bell (SRD)
2:00-3:30 Person Dups and EEs- Feldpausch (DSSD)

9/21 10:30-12:00 Preliminary Total Error Model and Loss Functions - Mulry (DSSD)

ESCAP MEETING NO. 67 - 09/17/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 67

September 17, 2001

Prepared by: Sarah Brady

The sixty-seventh meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 17, 2001 at 10:30. The agenda for the meeting was to discuss demographic analysis.

Committee Attendees:

Nancy Potok
John Thompson
Cynthia Clark
Nancy Gordon
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
Ruth Ann Killion
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines	Kevin Deardorff
Bill Bell	Lisa Blumerman
Tommy Wright	Fay Nash
Rita Petroni	Maria Urrutia
Raj Singh	Sarah Brady
Donna Kostanich	Kathleen Styles
Signe Wetrogan	Roxie Jones
Gregg Robinson	

I. Demographic Analysis

Overview of revisions made to DA since March

John Long began the presentation by giving an overview of the Demographic Analysis (DA) research. The Population Division (POP) has made several revisions to demographic analysis since March. At the August 1st ESCAP meeting (meeting #57), we discussed the revision made to the historical data and 1990 Population base and the revision to the birth component of the DA equation due to improved birth registration. John then turned the presentation over to Kevin Deardorff, who discussed the revisions made to the immigration components controlled to census results for change in the foreign-born population.

Evaluation of international migration

The first step in evaluating the international migration was to determine the foreign born population. On July 26, 2001 (meeting #55), the Committee discussed that the foreign born population was calculated from the Census 2000 long form data. The Census 2000 Supplementary Survey (C2SS) agreed with the Census 2000 long form on the foreign born population. POP then calculated each of the components of the foreign born population. These components are: legal population, emigrants, temporary (legal) migrants, and unauthorized migrants (residual migrants). The equation for the foreign born is:

$$FB_{1990\text{-}2000}=[L_{1990\text{-}2000}-(M+E)]+T_{1997\text{-}2000}+U_{1990\text{-}2000}$$

where:

FB=Census-based foreign born population
L=Legal population
M=Mortality
E=Emigrants
T=Temporary (legal) migrants
U=Unauthorized migrants

We examined each of the components of the foreign born equation, how they were obtained, and the limitations for determining the figures for each. We discussed the following:

Legal immigration– Includes people in the following categories: new arrival, people adjusting status to legal permanent resident, asylees, and refugees. The number of legal immigrants is calculated using data from INS. There are some legal immigrants that are not in this category because they couldn't be classified due to limitations. These legal immigrants are then included in the residual migrants component. There are several

limitations to the legal immigrants estimate:

- There are no data on nonimmigrants in the US who adjust legal status in the future. Thus, we assumed that future adjustments are similar to current adjustments.
- There is a backlog of applications for those who are adjusting their legal status.
- We also assume that these future adjustees have characteristics similar to current adjustees.

Emigrants– Includes the number of U.S. foreign born (naturalized citizens and permanent residents) who depart from the U.S. to reside abroad. Excludes unauthorized migrants, migrants from Puerto Rico, and temporary migrants. A residual methodology is used to obtain the annual number of emigrants and resulting emigration rate for the 1980-1990 decade by age, sex, race, Hispanic origin, and country of birth. The 1980-1990 foreign born emigration estimates are used for the 1990-2000 period. There are several limitations to the foreign born emigration methodology:

- We assumed complete coverage in both the 1980 and 1990 censuses.
- The application of trends from the previous decade was used to reflect trends within the most recent decade.
- We used aggregated race and Hispanic origin country groups to reflect trends for individual countries.
- A lengthy time interval was used to calculate estimates. This does not replicate actual fluctuations in trends occurring within the decade.

Temporary migrants– Includes those admitted to the United States for a specified purpose and temporary period but not for permanent residence– includes students and temporary workers; excludes tourists and business visitors. Using results from the C2SS for certain variables, we classified temporary migrants into types that correspond to VISA categories. There are some temporary migrants that are not classified in this component because of limitations. These temporary migrants are then included in the residual migrants component. There are several limitations to the methodology for the temporary migrant estimate:

- There is limited research on reasonableness of criteria used to identify temporary migrants, including income levels and occupations.
- There are recent categories of temporary migrants that were not specifically identified, including high tech specialty workers, treaty traders and investors, and North American Free Trade Agreement (NAFTA) workers.
- We also used 1990 group quarters proportions to estimate 2000 group quarters.

Unauthorized migrants (residual migrants)– Includes people who are illegally present in the United States. We assume that unauthorized migrants include the foreign born who were enumerated in the decennial census, and who were not otherwise accounted for in a legal migration component. The residual population is estimated by subtracting the other components of the foreign born population (legal, temporary, etc.) from the foreign born population. There are several limitations to the methodology for the residual migrant estimate:

- This component inherits the limitations of the other international migration components.
- The INS data are modified to be combined with census data.
- For the initial census-level calculations, we assumed 100 percent census coverage of foreign born regardless of legal status.
- We assumed 100 percent of Special Agricultural Workers were present in the U.S. on April 1, 1990.
- Some "humanitarian" populations (e.g., nonadjusted refugees/asylees) were omitted from the legal population, and, therefore, included in the residual migrant count.
- The Immigration Reform and Control Act (IRCA) legalized 1990 estimate was restricted to people granted permanent legal status; pending cases were treated as unauthorized in the 1990 estimate.
- The concept of "usual residence" and year of entry is unclear for migrants.
- There is a potential for misreporting citizenship status.
- We assumed there was no difference in race or Hispanic origin identification between administrative records and census.

POP then adjusted the unauthorized migrant component for a 15 percent undercount and for a 20 percent undercount. These adjustments were used to come up with two revised DA estimates. These estimates will be presented at the ESCAP meeting on September 20.

Howard Hogan raised several concerns about the evaluation of international migration and the assumptions for the evaluation. He will work with John Long and present his concerns and an alternative method to determine international migration at the next ESCAP meeting.

II. Next Meeting

The next meeting is scheduled for September 20, 2001 at 10:30. The agenda is to discuss the issues raised by Howard Hogan pertaining to demographic analysis.

ESCAP MEETING NO. 68 - 09/20/01

AGENDA

Kathleen P Porter
09/14/2001 02:09 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, RJones17@doc.gov, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, roxie.jones@mail.doc.gov, RJones17@doc.gov
cc: Donald H Keathley/PRED/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Mary Helen Mulry/SRD/HQ/BOC@BOC, Roxanne Feldpausch/DSSD/HQ/BOC@BOC
Subject: ESCAP Meetings for week of September 17

The following dates are scheduled ESCAP meetings for the week of September 17 (all are in Rm. 2412/3):

9/17 10:30-12:00 Remaining DA - POP

9/20 10:30-12:00 Missing Data - Keathley (PRED)
Correlation Bias - Bell (SRD)
2:00-3:30 Person Dups and EEs- Feldpausch (DSSD)

9/21 10:30-12:00 Preliminary Total Error Model and Loss Functions - Mulry (DSSD)

ESCAP MEETING NO. 68 - 09/20/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting #68

September 20, 2001

Prepared by: Nick Birnbaum

The sixty-eighth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 20, 2001 at 10:30 am. The agenda for the meeting was to discuss issues relating to the demographic analysis (DA) estimation of the foreign-born population.

Committee Attendees:

John Thompson
Ruth Ann Killion
Cynthia Clark
Jay Waite
Bob Fay
Howard Hogan
John Long
Carol Van Horn
Teresa Angueira
Nancy Potok
Nancy Gordon

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines	Donna Kostanich
Bill Bell	Nick Birnbaum
Kathleen Styles	Maria Urrutia
Sarah Brady	Raj Singh
Rita Petroni	Gregg Robinson
Fay Nash	Carolee Bush
Tommy Wright	Art Cresce
Kevin Deardorff	Lisa Blumerman
Signe Wetrogan	Roxie Jones

I. Assumptions Underlying the Demographic Estimation of the Foreign-Born Population

Howard Hogan discussed some of the assumptions underlying the demographic analysis estimation of the foreign born:

1) There is an assumption of no content error in the census. Content error includes both response error and error due to missing data. Content error could lead to an underestimate of the number of foreign born included in the census. That is, although included in the census, a given percentage may be misclassified or "imputed" as native born. Thus, the census count of the foreign born would underestimate the true number of foreign born included in the census (including those misclassified).

2) There is an assumption of complete coverage of the landed (authorized) immigrant and temporary worker populations in the census.

The implications of these assumptions are as follows:

- Higher levels of unauthorized immigration are necessary to bring the estimates closer in line with the A.C.E. results

- A specific age, sex, race, and origin composition of the foreign-born population missed by the census is implied.

With regard to assumption #1 above, some sizable percentage of the foreign born may answer the question on citizenship incorrectly, whether intentionally or not. Additionally, it may be the case that foreign-born households disproportionately do not respond to the long-form questionnaire, without regard to the neighborhood they live in. Thus, the Census Bureau's missing data models would tend to result in an underestimate of the number of foreign born included in the census.

With regard to assumption #2 above, it requires that the foreign born missed in the census must all come from the unauthorized immigration component. Thus, the unauthorized component must "absorb" the entire undercount for the foreign-born population. The demographic analysis equation, by assuming complete coverage for the landed immigrant component, might produce an estimate of the number of unauthorized immigrants included in the census that is too low. For any assumed "reasonable" undercount rate for this population, the estimate should be higher.

This inquiry raises the question as to how much larger the foreign-born population would be under this alternative methodology than that currently derived by DA estimation, and how the age, sex, race, and origin composition would differ.

John Long provided a brief summary of the demographic analysis perspective on some of the issues raised by Howard in his presentation. These points will be developed in further detail in a future Population Division staff presentation to the Committee.

John noted that content error (assumption #1) is extremely hard to measure and there is little evidence that points to the direction of that error. Howard has given the reasons that it might underestimate the foreign-born population, but there are also reasons to expect overestimation of the foreign born. As an example, if the Census Bureau did not obtain a response to the citizenship question for the children in a household where the parents were foreign born, the children would always be imputed as foreign born, whereas we clearly know from other sources that more than a trivial number of these children are not foreign born. Additionally, estimates of the foreign born from Census 2000 and the Census 2000 Supplemental Survey are consistent.

Investigation of the effects of changing assumption #2 is underway and will be presented at a later ESCAP presentation.

In closing the meeting, John Thompson summarized the main areas of uncertainty with regard to the size and composition of the estimate of the foreign-born population:

- The level of census under-coverage of the legal components of the foreign-born population

- Whether content error in the census results in the census counts of the foreign born understating the number included in the census, regardless of classification as authorized or unauthorized.

- The level of census under-coverage of the unauthorized component of the foreign-born population.

II. Next Meeting

The agenda for the next meeting, scheduled for the afternoon of September 20, 2001, is to discuss: 1) the revised demographic analysis estimates, and 2) the alternative missing data models.

ESCAP MEETING NO. 69 - 09/20/01

AGENDA

Kathleen P Porter
09/14/2001 02:09 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, RJones17@doc.gov, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, roxie.jones@mail.doc.gov, RJones17@doc.gov
cc: Donald H Keathley/PRED/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Mary Helen Mulry/SRD/HQ/BOC@BOC, Roxanne Feldpausch/DSSD/HQ/BOC@BOC
Subject: ESCAP Meetings for week of September 17

The following dates are scheduled ESCAP meetings for the week of September 17 (all are in Rm. 2412/3):

9/17 10:30-12:00 Remaining DA - POP

9/20 10:30-12:00 Missing Data - Keathley (PRED)
Correlation Bias - Bell (SRD)
2:00-3:30 Person Dups and EEs- Feldpausch (DSSD)

9/21 10:30-12:00 Preliminary Total Error Model and Loss Functions - Mulry (DSSD)

ESCAP MEETING NO. 69 - 09/20/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 69

September 20, 2001

Prepared by: Sarah Brady

The sixty-ninth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 20, 2001 at 2:00. The agenda for the meeting was to discuss the revised demographic analysis estimates and alternatives models for missing data.

Committee Attendees:

Nancy Potok
John Thompson
Cynthia Clark
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
Ruth Ann Killion
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines	Kevin Deardorff
Bill Bell	Lisa Blumerman
Dave Hubble	Fay Nash
Rita Petroni	Maria Urrutia
Don Keathley	Sarah Brady
Anne Kearney	Kathleen Styles
Donna Kostanich	Roxie Jones
Signe Wetrogan	Carolee Bush
Gregg Robinson	Pat Cantwell

I. Revised Demographic Analysis Estimates

Gregg Robinson presented the revised demographic analysis (DA) estimates. These estimates included all of the revisions to DA discussed at ESCAP meetings #55 on July 26, 2001, #57 on August 1, 2001, and #67 on September 17, 2001. Two revised estimates were presented. One estimate assumed an undercount of 15 percent for the residual immigrants component; the other assumed an undercount of 20 percent.

- The revised DA estimate of the population assuming 15 percent undercount of the residual population is 281,759,875. The revised DA estimate assuming 20 percent undercount of the residual population is 282,399,979. These estimates surround the alternative DA estimate of 282,335,711.

Gregg Robinson also presented the revised DA estimates by race (Black/Nonblack), sex, and age. The net undercounts implied by these estimates were compared to the A.C.E. estimates of undercounts for these variables. The following were notable findings:

- The three DA estimates imply a percent net overcount for Nonblacks, while A.C.E. measured a percent net undercount for this group.

- The three DA estimates imply a percent net overcount for females, while A.C.E. measured a percent net undercount for this group.

- Although the three DA estimates and the A.C.E. imply percent net undercounts for males ages 0-17 and 18-29, the percent net undercount measured by A.C.E. is much larger than that measured by the DA estimates.

- The percent net undercounts measured by the DA estimates and A.C.E. for males age 30-49 are similar.

- The largest difference in the percent net undercount for females is in the 18-29 age category. The net undercount for A.C.E. is 2.11 percent; the net undercount for the DA estimates ranges from -0.66 percent to -1.74 percent.

Staff from the Population Division will now look at scenarios that will move the DA estimates to the level of the A.C.E. estimates. They will present these data at a meeting next week. The Committee will examine the data and whether the assumptions behind the scenarios are plausible.

II. Alternative Models of Missing Data

Alternative models of missing data were discussed at ESCAP meeting #59, on August 16, 2001. Several concerns were expressed by Committee members about the evaluation. An interdivisional team, called the missing data alternative planning commission,

conducted additional research to address the Committee's concerns. This meeting was held to discuss the results of the team's research.

As discussed at the previous meeting, A.C.E. used the following missing data procedures: a noninterview adjustment; characteristic imputation for race, ethnicity, tenure, sex, and age; and probability imputation for enumeration, match, and resident statuses. Several alternatives were discussed at the meeting. These original alternatives are described in the minutes from ESCAP meeting #59.

The additional research grouped the original alternatives into four groups. These groups are as follows:

Group 1– Included alternative noninterview adjustment cells, nearest neighbor imputation for noninterview adjustment, logistic regression for the probability imputation, late census data, and no non-ignorable missingness for probability imputation. There were 16 alternatives in this group.

Group 2– Included alternative noninterview adjustment cells, nearest neighbor imputation for noninterview adjustment, logistic regression for the probability imputation, late census data, and non-ignorable missingness for all three probability imputations. There were 16 alternatives in this group.

Group 3– Included alternative noninterview adjustment cells, nearest neighbor imputation for noninterview adjustment, logistic regression for the probability imputation, non-ignorable missingness for either one or two probability imputations, and no late census data combinations. There were 48 alternatives in this group.

Group 4– Included alternative noninterview adjustment cells, nearest neighbor imputation for noninterview adjustment, logistic regression for the probability imputation, late census data, and non-ignorable missingness for either one or two probability imputations, and late census data combinations only. There were 48 alternatives in this group.

Don Keathley presented the results on the Dual System Estimates (DSEs) from running the missing data system on each of these four groups.

- The range of DSEs between groups 1 and 2 were similar (1,266,317.34 and 1,300,959.23, respectively).

- The range of the DSEs for group 3 was 1,750,773.05. This is 484,455.71 and 449,813.72 larger than the ranges for groups 1 and 2.

- The range of the DSEs for group 4 was 2,628,487.66, which was the largest range among the groups.

- Alternative noninterview adjustment cell definitions increased the DSEs, except when combined with both late census data and logistic regression. They also produced the highest DSEs when combined with both nearest neighbor imputation for noninterview adjustment and late census data.

- Both the nearest neighbor imputation for noninterview adjustment and late census data had no apparent effect on the DSEs.

- Logistic regression for probability imputation, non-ignorable missingness for enumeration status, and non-ignorable missingness for resident status decreased the DSEs.

- Non-ignorable missingness for match status increased the DSEs.

- The tandem of late data and logistic regression decreased the DSEs and resulted in the lowest DSEs when taken by themselves.

The next step involving missing data will be to incorporate its effect into the total error model as a random or variance component.

III. Next Meeting

The next meeting is scheduled for September 21, 2001 at 9:00. The agenda is to discuss the preliminary total error model and loss functions.

ESCAP MEETING NO. 70 - 09/21/01

AGENDA

Kathleen P Porter
09/14/2001 02:09 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, RJones17@doc.gov, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, roxie.jones@mail.doc.gov, RJones17@doc.gov
cc: Donald H Keathley/PRED/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Mary Helen Mulry/SRD/HQ/BOC@BOC, Roxanne Feldpausch/DSSD/HQ/BOC@BOC
Subject: ESCAP Meetings for week of September 17

The following dates are scheduled ESCAP meetings for the week of September 17 (all are in Rm. 2412/3):

9/17 10:30-12:00 Remaining DA - POP

9/20 10:30-12:00 Missing Data - Keathley (PRED)
Correlation Bias - Bell (SRD)
2:00-3:30 Person Dups and EEs- Feldpausch (DSSD)

9/21 10:30-12:00 Preliminary Total Error Model and Loss Functions - Mulry (DSSD)

ESCAP MEETING NO. 70 - 09/21/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting #70

September 21, 2001

Prepared by: Nick Birnbaum

The seventieth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 21, 2001 at 9:00 am. The agenda for the meeting was a preliminary examination of the revised total error model using component values from the A.C.E. evaluation data, and a discussion of loss function analysis.

Committee Attendees:

John Thompson
Ruth Ann Killion
Cynthia Clark
Jay Waite
Bob Fay
Howard Hogan
John Long
Teresa Angueira
Nancy Potok
Nancy Gordon

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines
Dave Hubble
Kathleen Styles
Sarah Brady
Fay Nash
Donna Kostanich
Nick Birnbaum
Maria Urrutia
Rita Petroni
Carolee Bush
Mary Mulry (via video conferencing)

I. Preliminary Examination of Revised Total Error Model Using Component Values from the A.C.E. Evaluation Data and Discussion of Loss Function Analysis

Total Error Model

The total error results from the revised total error model (TEM) were presented. It should be noted that the results presented are preliminary in nature and are subject to the following limitations, among others:

- The P-sample data collection error, P-sample fabrication, and the E-sample data collection error will be revised, based on recent reviews of the Evaluation Followup data.

- The correlation bias estimates used here are based on the March 2001 "alternative" demographic analysis (DA) estimates. Revised DA estimates from September 2001 will be used as the basis for the correlation bias estimates in the final total error model.

- The values for imputation error used in this analysis were the 1990 reasonable alternative imputations adjusted for 2000. The final analysis will use the values from the 2000 reasonable alternative imputations model.

The results obtained from the preliminary total error model were reviewed. The most visible difference in the error components is that the E-sample data collection error is much larger than the 1990 value used in February. However, this increase was offset by decreases in other error components, including E-sample processing error and P-sample data collection and matching errors.

Questions were raised about the sources of the values for the error components in the total error model and about what seemed to be inconsistencies in the results. Committee members requested specific information about how the data presented from the various evaluations and studies related to the values for the error components in the preliminary TEM. Additional information was required to validate that the total error model was correctly including the results of the individual evaluation studies. It was agreed that this would have to be examined in greater detail outside of the meeting.

In the total error model, the accuracy of the census counts are compared to the A.C.E. estimates using confidence intervals for the "true" undercount based on both bias and variance in the A.C.E. obtained from the error components. The comparisons were carried out for the sixteen evaluation post-strata and for the nation as a whole.

Loss Function Analysis

It was noted that the inconsistencies in the TEM must be resolved before the loss function analyses can be studied.

The loss function analysis compares the accuracy of the levels and shares of the census counts and the A.C.E. estimates for various geographic areas or groupings. The results of the total error simulations produce bias-corrected dual system estimates that are used as estimates of the target population in the loss function analysis, which also takes into account the variance in the estimates of the targets. Two different methods are used to allocate the bias to the A.C.E. post-strata. The same four assumptions regarding correlation bias in the total error model are used here.

Whether one focuses on levels or shares and at what geographic level depends upon the uses of the data. Since the pending decision relates to funds allocation and survey controls, both levels and shares are important for various areas or groupings.

The following table shows the geographic areas or groupings examined, and the corresponding loss measurements associated with the specified uses of the data:

Areas/Groups	**Loss Measurement**	**Uses of the Data**
All counties	Shares within states	Funds allocation within state
All counties w/pop. greater than 65,000	Levels	Areas for American Community Survey (ACS) estimates
All counties w/pop. greater than 100,000	Levels	Federal funds allocation threshold
All counties w/pop. greater than 65,000 by demographic group	Levels	ACS weighting cells
All places	Shares within states	Funds allocation within state
All places w/pop. between 25,000 and 50,000	Levels and shares within the US	Federal funds allocation
All places w/pop. between 50,000 and 100,000	Levels and shares within the US	Federal funds allocation
All places w/pop. greater than 100,000	Levels and shares within the US	Federal funds allocation
All places w/pop. greater than 50,000	Levels	Federal funds allocation threshold
States	Levels and shares within the US	Federal funds allocation and Current Population Survey weighting cells

Subsequent to the meeting, John Thompson asked Rita Petroni to calculate additional loss functions (levels) for: counties with a population of less than 100,000; counties with a population of less than 65,000; and places with a population of less than 50,000.

The final total error model and loss function analysis results will not be available until early October. This is because, as mentioned above, some of the component data are not yet available, including new estimates of correlation bias, or are currently being revised. It is assumed that the new correlation bias estimates to be used in the final total error model and loss functions won't change substantially from the earlier ones.

II. Next Meeting

The agenda for the next meeting, scheduled for September 25, 2001, is to discuss the revised estimates of correlation bias in the A.C.E. estimates.

ESCAP MEETING NO. 71 - 09/25/01

AGENDA

Kathleen P Porter
09/20/2001 02:04 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, RJones17@doc.gov, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, roxie.jones@mail.doc.gov, RJones17@doc.gov, Jacqueline M Cusick/DIR/HQ/BOC@BOC

cc: Bonnie J Demarr/DSSD/HQ/BOC@BOC, Elizabeth A Krejsa/PRED/HQ/BOC@BOC, Tamara S Adams/DSSD/HQ/BOC@BOC, Mary Helen Mulry/SRD/HQ/BOC@BOC, Roxanne Feldpausch/DSSD/HQ/BOC@BOC

Subject: ESCAP Meetings for week of September 24

The ESCAP Meetings for the week of September 24 are as follows (all are in Rm. 2412/3):

September 25 10:30-12:00 Correlation Bias - Bell (SRD)

September 26 10:30-12:00 Evaluation and Person Followup Questionnaires - Martin (DIR)

EFU Rework - Adams/Krejsa (DSSD/PRED)

Person Dups and EEs - Feldpausch (DSSD)

September 28 1:30-3:00 Final Total Error and Loss Functions - Mulry

ESCAP MEETING NO. 71 - 09/25/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 71

September 25, 2001

Prepared by: Sarah Brady

The seventy-first meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 25, 2001 at 10:30. The agenda for the meeting was to discuss correlation bias.

Committee Attendees:

Nancy Potok
John Thompson
Cynthia Clark
Nancy Gordon
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
Ruth Ann Killion
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Bill Bell	Maria Urrutia
Tommy Wright	Sarah Brady
Rita Petroni	Kathleen Styles
Raj Singh	Nick Birnbaum
Fay Nash	

I. Briefing on the presentation of demographic analysis

The Population Division (POP) met this week with several outside expert demographers to discuss their analysis of the international migration that is now included in the revised demographic analysis (DA) estimates. The group was principally asked to comment on the temporary migration, emigration, and unauthorized migration components. The INS has information that could be useful for the temporary migration category, resulting in an increase of approximately 418,000 to this population. This information will not be incorporated into the revised DA for the purposes of ESCAP, but will be incorporated at a later date. It was suggested that we should refer to the unauthorized migration component as the residual migration, since it contains other types of foreign born that could not be classified into any other component and therefore were placed with the unauthorized.

II. Correlation bias

Correlation bias in Dual System Estimates (DSEs) results from a failure of the general independence assumption underlying DSEs due to either causal dependence or heterogeneity. Causal dependence occurs when the act of being included in the census makes someone more or less likely to be included in the A.C.E. Heterogeneity occurs when the census and A.C.E. inclusion probabilities vary over persons within post-strata. When heterogeneity within post-strata exists, it is generally suspected to be of the form where persons more likely to be missed in the census are also more likely to be missed in the A.C.E. This will lead to an underestimation of the true population by the DSEs.

Correlation bias is calculated only for adult males; the method assumes there is no correlation bias for adult females or children. Correlation bias is estimated via several alternative models using the DA sex ratios and A.C.E. data. The correlation bias estimates available for the March ESCAP recommendation used DA estimates as of February 16, 2001.

Bill Bell presented new correlation bias estimates using the alternative DA estimates from February 26, 2001 and also using the current (September 2001) revised DA estimates. The table below illustrates the percent correlation bias estimates for 2000 (using the three DA estimates). Results are also shown for 1990.

Percent Correlation Bias Estimates (two-group model) for 2000 A.C.E. (Using alternative DA estimates) and 1990 PES

	Original DA 2/16/01	Alt. DA March 2001	Revised DA 15% undercount	1990 DA
Black				
18-29 M	-7.37	-7.30	-6.91	-8.01
30-49 M	-8.10	-7.93	-8.26	-7.70
50+ M	-4.74	-4.60	-4.95	-8.22
Nonblack				
18-29 M	2.47	1.69	0.41	-0.32
30-49 M	-0.45	-0.56	-0.85	-1.64
50+ M	-0.74	-0.74	-0.79	-1.17

Although the estimate of correlation bias for nonblack males age 18-29 has moved closer to zero, a positive value (indicating DSE overestimation) is still unexpected. We would expect the correlation bias to be negative. Howard Hogan proposed three hypotheses (meant to be exhaustive) as to why the measured estimate of percent correlation bias for this category is positive: 1) there was more correlation bias for females than males; 2) the processing errors associated with A.C.E. are differential– there were more errors for males than females; 3) the residual immigration estimates for this category are wrong. The first hypothesis seems unlikely; the second and third relate to data errors that would suggest the data cannot support estimation of correlation bias for this group.

The estimates of correlation bias for the revised DA estimates will be used in the total error model. Correlation bias for nonblack males age 18-29 will be set to zero.

III. Next Meeting

The next meeting is scheduled for September 26, 2001 at 10:30. The agenda is to discuss the rework done for the Evaluation Followup and Person Followup.

ESCAP MEETING NO. 72 - 09/26/01

AGENDA

Kathleen P Porter
09/20/2001 02:04 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, RJones17@doc.gov, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, roxie.jones@mail.doc.gov, RJones17@doc.gov, Jacqueline M Cusick/DIR/HQ/BOC@BOC
cc: Bonnie J Demarr/DSSD/HQ/BOC@BOC, Elizabeth A Krejsa/PRED/HQ/BOC@BOC, Tamara S Adams/DSSD/HQ/BOC@BOC, Mary Helen Mulry/SRD/HQ/BOC@BOC, Roxanne Feldpausch/DSSD/HQ/BOC@BOC
Subject: ESCAP Meetings for week of September 24

The ESCAP Meetings for the week of September 24 are as follows (all are in Rm. 2412/3):

September 25 10:30-12:00 Correlation Bias - Bell (SRD)

September 26 10:30-12:00 Evaluation and Person Followup Questionnaires - Martin (DIR)
EFU Rework - Adams/Krejsa (DSSD/PRED)
Person Dups and EEs - Feldpausch (DSSD)
September 28 1:30-3:00 Final Total Error and Loss Functions - Mulry

ESCAP MEETING NO. 72 - 09/26/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting #72

September 26, 2001

Prepared by: Nick Birnbaum

The seventy-second meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 26, 2001 at 10:30 am. The agenda for the meeting was to discuss the results of the Person Followup (PFU) and Evaluation Followup (EFU) forms review.

Committee Attendees:

John Thompson
Ruth Ann Killion
Cynthia Clark
Jay Waite
Bob Fay
Howard Hogan
John Long
Teresa Angueira
Nancy Potok
Nancy Gordon
Carol Van Horn

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines	Bill Bell
Betsy Martin	Nick Birnbaum
Kathleen Styles	Maria Urrutia
Sarah Brady	Rita Petroni
Fay Nash	Raj Singh
Elizabeth Krejsa	David Raglin
Tommy Wright	Tammy Adams
Tom Mule	Debbie Fenstermaker
Danny Childers	Roxanne Feldpausch
David Whitford	

I. Results of the Person Followup and Evaluation Followup Forms Review

On July 27, 2001, at ESCAP meeting #56, the Committee was presented with data comparing production match codes (includes cases that went to PFU and cases that were coded based on the initial interview only) to Measurement Error Reinterview (MER or Evaluation Followup (EFU)) codes from the *Analysis of Measurement Error Study*. The EFU was a review of production cases – both matches from the initial interviewing and person followup cases for the non-matches. After reviewing the results from the July 27 presentation, the Committee determined that an additional review of a sample of the production and EFU forms for the E-sample was warranted to determine if the proper codes has been assigned on a consistent basis and examine cases where production and the EFU produced conflicting codes.

This review was an analyst-only operation; that is, only the most highly trained and experienced personnel who work on these operations were used. Each analyst reviewed a work unit of sampled persons, coding the EFU form independently of the production cases. The goal was, to the extent possible, to resolve the concerns identified in the MER. An important aspect to the coding procedures was the consistent application of the residence rules.

After coding each form independently, the analyst would indicate which contained the correct code – both, EFU, production, or conflicting. Cases where a clear determination of the correct code could not be made were deemed conflicting; either contradictory information was provided by the same respondent type (both non-proxies or similar proxies) or the geocoding information for the housing unit was contradictory.

One of the results from the presentation in July was that the MER showed a net difference in erroneous enumeration coding of approximately 1.9 million; that is, there were
1.9 million more production correct enumerations that were coded as erroneous enumerations in the MER than vice versa. The unresolved rate for the MER coding was 1.7%, whereas the unresolved rate for the production cases in the MER sample was 2.6%. The PFU/EFU review determined the "best" code and compared it to production to see how much of a change there would be in the net difference in erroneous enumeration coding. This difference dropped from 1.9 million to 1.45 million. The number of production correct enumerations with a PFU/EFU review best code of erroneous enumeration fell from 2,827,414 to 1,816,315. Of this latter number, 62.7% of these were production matched cases. The number of production erroneous enumerations with a PFU/EFU review best code of correct enumeration fell from 908,385 to 361,400. However, the percent unresolved increased to 4.82% and conflicting cases constituted 0.99%. The causes of the changes from correct to erroneous enumeration coding between the production and the "best" code in the PFU/EFU review included coding error, conflicting cases (not allowed in the MER – at least some of these cases would

have been coded as erroneous enumerations in that evaluation), and an increased unresolved rate.

The review provided data on the reasons cases coded as correct enumerations in production were coded as erroneous enumerations according to the best code from this additional review and vice versa. For example, approximately 57% of the erroneous enumerations missed by production were now reported either at a group quarters or at a second home. Similar data were provided on EFU erroneous enumerations that were coded as "best non-erroneous" (correct or unresolved) in the PFU/EFU review.

The review indicated a production coding error rate of 0.68% (comparable to the Matching Error Study E-sample gross error rate of 0.62%) and an MER coding error rate of approximately 3.4%. This review of the EFU cases resulted in an increase in the unresolved rate to 9.38%.

Data were also presented on the PFU/EFU review cases that were coded as unresolved. Most of these, from both the production and EFU, were coded unresolved due to a residence rules issue. Data on the conflicting cases – the production review indicated an erroneous enumeration while the EFU review indicated a correct enumeration or vice versa – showed that those with contradictory geocoding information or involving movers were the majority of these cases.

Finally, proxy respondent data from the review samples of the two forms were presented. In the case of non-proxy interviews for both forms (EFU and PFU), the codes from both forms were selected approximately 87% of the time. If both were not selected, then, as expected, the code from the form with the non-proxy interview was chosen more often.

In discussing the results of the PFU/EFU review, it was noted that the production correct enumeration rate (does not include insufficient information and duplicate cases) for the sample was 97.77%. However, the estimate of the overall correct enumeration rate in the A.C.E. would depend upon the correct enumeration probabilities for the unresolved and conflicting cases. Thus, depending upon one's assumptions regarding the correct enumeration probabilities for these cases, these results suggest the potential for a non-trivial effect on the correction enumeration rate and consequently on the dual system estimate itself.

II. Next Meeting

The agenda for the next meeting, scheduled for September 27, 2001, is to discuss the results of: 1) the EFU and PFU questionnaire study, and 2) the analysis of census person duplicates and the corresponding A.C.E. enumeration code.

ESCAP MEETING NO. 73 - 09/27/01

AGENDA

Kathleen P Porter
09/20/2001 02:04 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, RJones17@doc.gov, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, roxie.jones@mail.doc.gov, RJones17@doc.gov, Jacqueline M Cusick/DIR/HQ/BOC@BOC

cc: Bonnie J Demarr/DSSD/HQ/BOC@BOC, Elizabeth A Krejsa/PRED/HQ/BOC@BOC, Tamara S Adams/DSSD/HQ/BOC@BOC, Mary Helen Mulry/SRD/HQ/BOC@BOC, Roxanne Feldpausch/DSSD/HQ/BOC@BOC

Subject: ESCAP Meetings for week of September 24

The ESCAP Meetings for the week of September 24 are as follows (all are in Rm. 2412/3):

September 25 10:30-12:00 Correlation Bias - Bell (SRD)

September 26 10:30-12:00 Evaluation and Person Followup Questionnaires - Martin (DIR)

EFU Rework - Adams/Krejsa (DSSD/PRED)

Person Dups and EEs - Feldpausch (DSSD)

September 28 1:30-3:00 Final Total Error and Loss Functions - Mulry

ESCAP MEETING NO. 73 - 09/27/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 73

September 27, 2001

Prepared by: Sarah Brady

The seventy-third meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 27, 2001 at 10:30. The agenda for the meeting was to discuss a review of the Person Followup (PFU) and Evaluation Followup (EFU) forms and the evaluation of census duplicates and the corresponding A.C.E. coding.

Committee Attendees:

John Thompson
Cynthia Clark
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
Ruth Ann Killion
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Bill Bell	David Whitford
Marvin Raines	Danny Childers
Rita Petroni	Roxanne Feldpausch
Dave Raglin	Betsy Martin
Elizabeth Krejsa	Maria Urrutia
Raj Singh	Sarah Brady
Debbie Fenstermaker	Kathleen Styles
Tom Mule	Roxie Jones

I. Review of the PFU and EFU instruments

As stated at the previous ESCAP meeting on September 26, approximately 1.8M production correct enumerations were classified as erroneous enumerations during the PFU/EFU review. These results included cases that were matched during production and were not sent to PFU. The Committee felt that it was important to examine the followup instruments to see if the differences between them could explain the classification change for the cases that went out to both PFU and EFU. Betsy Martin reviewed the instruments and relevant research and evidence to assess the probable effect on the estimates of erroneous enumerations each instrument would have for major coverage categories. The coverage categories were apartment mixups, movers, group quarters residents, college students in dorms, college students in other housing units, assisted living or group care facilities, and multiple residences.

Betsy identified positive and negative instrument features pertaining to each category and what would be the predicted effect. She then looked at what this effect would do to the estimates of erroneous enumerations. The net projected effect for EFU is that it would slightly overestimate erroneous enumerations, especially in situations of apartment mixups, moves within the search area, people in assisted living housing units and college students in housing units, due to the fact that it did not always obtain addresses for alternative addresses. (EFU may also be characterized by other, unknown biases.) The net projected effect for PFU is that it would considerably underestimate erroneous enumerations, especially in situations of college students in dorms, other group quarters, and other residences. Although EFU may slightly overestimate erroneous enumerations, she felt it is probably closer to the truth.

II. Census person duplication and the corresponding A.C.E. enumeration status

The Committee reviewed data from the Evaluation of Census Person Duplication at the ESCAP meeting (#65) on September 12, 2001. As discussed at that meeting, we conducted another evaluation which examined the A.C.E. enumeration codes for the people identified as duplicates outside the A.C.E. clusters. In doing this, we could determine if the A.C.E. correctly measured duplication in the census as erroneous enumerations. Therefore, this could also give the Committee insight into the unresolved and conflicting cases from the Person Followup (PFU) and Evaluation Followup (EFU) forms review. For more details about this evaluation, see the minutes from ESCAP meeting #72 on September 26, 2001.

Roxanne Feldpausch presented results from the evaluation of census person duplication and the corresponding A.C.E. enumeration status. The results were as follows:

- The percent of erroneous enumerations for the E-Sample people duplicated to people in housing units outside the A.C.E. search area was 14.2 percent; this was lower than what we expected. We would expect about 50 percent of the E-

Sample duplicates to people outside the A.C.E. search area to be erroneous enumerations. We would expect 50 percent, because half of the time the wrong housing unit should be in sample, resulting in coding the residents as erroneous.

- The percent erroneous enumeration for E-Sample people duplicated to people in group quarters when residents were not allowed to claim usual home elsewhere was 45.5 percent for college dorms and 16.5 percent for other group quarters; this was lower than what we expected.

Given the complexity of this evaluation, further analysis is needed to fully understand the implications of this evaluation. John Thompson tasked Bob Fay with presenting additional data on the relationship of erroneous enumerations to duplicates on October 1.

III. Next Meeting

The next meeting is scheduled for September 28, 2001 at 1:30. The agenda is to discuss demographic analysis (DA) scenarios that would bring the DA in line with the A.C.E.

ESCAP MEETING NO. 74 - 09/28/01

AGENDA

Kathleen P Porter

09/20/2001 02:04 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, RJones17@doc.gov, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, roxie.jones@mail.doc.gov, RJones17@doc.gov, Jacqueline M Cusick/DIR/HQ/BOC@BOC cc: Bonnie J Demarr/DSSD/HQ/BOC@BOC, Elizabeth A Krejsa/PRED/HQ/BOC@BOC, Tamara S Adams/DSSD/HQ/BOC@BOC, Mary Helen Mulry/SRD/HQ/BOC@BOC, Roxanne Feldpausch/DSSD/HQ/BOC@BOC

Subject: ESCAP Meetings for week of September 24

The ESCAP Meetings for the week of September 24 are as follows (all are in Rm. 2412/3):

September 25 10:30-12:00 Correlation Bias - Bell (SRD)

September 26 10:30-12:00 Evaluation and Person Followup Questionnaires - Martin (DIR)
EFU Rework - Adams/Krejsa (DSSD/PRED)
Person Dups and EEs - Feldpausch (DSSD)

September 28 1:30-3:00 Final Total Error and Loss Functions - Mulry

[**Note**: the topic for the September 28 meeting was changed to "Alternative DA Scenarios." A revised agenda for meeting #74 was not distributed.]

ESCAP MEETING NO. 74 - 09/28/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting #74

September 28, 2001

Prepared by: Nick Birnbaum

The seventy-fourth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on September 28, 2001 at 1:30 pm. The agenda for the meeting was to discuss alternative assumptions for the components of international migration and alternative demographic analysis (DA) scenarios.

Committee Attendees:

John Thompson
Ruth Ann Killion
Cynthia Clark
Jay Waite
Bob Fay
Howard Hogan
John Long
Teresa Angueira
Nancy Potok
Nancy Gordon
Carol Van Horn

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines	Bill Bell
Gregg Robinson	David Hubble
Art Cresce	Kevin Deardorff
Lisa Blumerman	Nick Birnbaum
Kathleen Styles	Maria Urrutia
Carolee Bush	Rita Petroni
Raj Singh	Tommy Wright

I. Alternative Assumptions for the Components of International Migration

Population Division staff presented data relating to alternative assumptions for the components of international migration. This discussion was a follow-up to the earlier ESCAP presentation on September 20, 2001 (meeting #68) entitled "Assumptions Underlying the Demographic Estimation of the Foreign-Born Population."

In estimating the foreign-born population, it was necessary to use Census 2000 data on the foreign-born population. Concerns were raised at the September 20, 2001 meeting about using these data to estimate the foreign-born population. Among the concerns were the following:

- There may be inconsistent reporting in response to the citizenship question on the part of the foreign born

- There may be higher levels of non-response to the long-form questionnaire among the foreign born, as well as higher levels of item non-response to the citizenship question

- The demographic estimate may not take into account the census under-coverage of the legal immigrant and temporary worker populations. A collateral concern is the assumed undercount rate of the residual foreign born.

The purpose of this presentation, summarized below, was to address these concerns.

The census data on foreign born come from the citizenship question on the long form. Native is defined as born in the US, or born in Puerto Rico or the Island Areas, or born overseas of American parents. Foreign born includes both non-citizens and naturalized citizens.

While there may be some non-trivial percentage of the foreign born that answers the question on citizenship incorrectly, data from 1990 census studies indicate that citizenship was reported consistently, if not altogether accurately. In fact, the studies indicate that citizenship was reported more consistently than other "sensitive" items including race, ancestry, and educational attainment. There was some differentiation in the level of consistency within population subgroups, but the results indicated consistent reporting across all groups.

With regard to missing data for this item, weighted preliminary Census 2000 sample data indicate that allocations were conducted approximately 5% of the time. This is in addition to the 10% of long forms that were not sample-defined. Among those reporting citizenship, 89.3% reported native and 10.7% reported foreign born. However, because there was greater non-response among the foreign born, the allocated percentage of foreign born was higher – 17.5% foreign born versus 82.5% native.

The Census 2000 data were also compared to the Census 2000 Supplementary Survey (C2SS) data on foreign born. Because the C2SS included item nonresponse followup on a sample basis, the Census Bureau was able to learn more about the non-responding population (unlike in the census). There was no evidence of a significant downward bias in the estimate of the foreign born from the C2SS enumeration or editing procedures. In fact, the C2SS foreign born weighted total is nearly identical to the preliminary Census 2000 sample data weighted total for households: 30,555,510 for Census 2000 versus 30,523,176 for C2SS.

The presentation then included a summary of recent discussions with international migration experts to obtain their input regarding the methodology and assumptions used in estimating the international migration components of the demographic analysis estimates. One of the suggestions that came out of this discussion was to estimate the unauthorized migrant population by removing known components from the residual foreign born. That is, there are known components of the population formerly referred to as the unauthorized migrant population. These components are: refugees and asylee applicants whose cases have not been processed yet because of INS backlogs, deported migrants, and the illegal population that legalized during the decade. The estimate of the total for the known components is 1.7 million. The balance of the residual foreign born is the (implied) unauthorized migrant population.

To address the concerns regarding the assumed undercount rates of both the legal and unauthorized components of the foreign born raised earlier and to incorporate the beneficial guidance provided by the external experts, several different estimates for the total foreign-born population were then presented to the Committee, based on differing assumptions regarding undercount rates for the various components. All of these estimates used 1.7 million as the "base" estimate for the known components of the residual foreign born (see discussion above); however, some of the estimates adjusted this component for an assumed undercount. Among the estimates for the total foreign-born population were: the census level estimate of 31,098,945 (no assumed undercount); an estimate of 32,635,199 based on the assumption of a 15% undercount for the residual foreign born; and an estimate of 33,091,988 based on the assumptions of a 5% undercount for the known components of the residual foreign born and a 12.5% undercount for the (implied) unauthorized migrant population.

II. Alternative Demographic Analysis Scenarios

Finally, Population Division staff used the revised DA estimate (presented on 20, 2001 (meeting #69) to the Committee – assumes a 15 percent undercount for the residual foreign born) and two alternative scenarios regarding the international migration components to demonstrate the implied age, race (Black/non-Black), and sex distributions for the various DA estimates. The alternative scenarios used were: (1) increasing the international migration components using the distribution of the residual foreign born so that the total DA estimate matches the A.C.E. population level, and (2) increasing the international migration components using the distribution of the total foreign born so that the total DA estimate matches the A.C.E. population level.

The purpose of the analysis was to see the implied age, race, and sex distributions when we try to close the gap between the total DA estimate and the A.C.E. population level. For example, scenario (1) would require an increase in Blacks among the residual foreign born of approximately 386,000 and an increase in females among the residual foreign born of approximately 1.34 million. However, given the known biases in the A.C.E., it was suggested that perhaps a more appropriate analysis would be to vary the international migration component estimates to produce total DA estimates that match a *bias-corrected* A.C.E. One could then examine the implied age, race, and sex distributions for those DA estimates.

III. Next Meeting

The agenda for the next meeting, scheduled for October 1, 2001, is to discuss:
1) remaining issues for the Committee to address, and 2) an analysis of the relationship between duplicates and erroneous enumerations in Census 2000.

ESCAP MEETING NO. 75 - 10/01/01

AGENDA

Kathleen P Porter
09/28/2001 03:37 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/EPCD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, Jacqueline M Cusick/DIR/HQ/BOC@BOC, roxie.jones@mail.doc.gov
cc:
Subject: Monday ESCAP Meeting

The agenda for the October 1 ESCAP meeting scheduled from 9-12 in Rm. 2412/3 is as follows:

John Thompson - Issues

Bob Fay - Duplicate Analysis

ESCAP MEETING NO. 75 - 10/01/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 75

October 1, 2001

Prepared by: Sarah Brady

The seventy-fifth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on October 1, 2001 at 9:00. The agenda for the meeting was to discuss issues the Committee must resolve in order to reach a recommendation and the relationship between duplicates and erroneous enumerations.

Committee Attendees:

Nancy Potok
John Thompson
Nancy Gordon
Cynthia Clark
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Bill Bell

Marvin Raines

Raj Singh

Maria Urrutia

Sarah Brady

Kathleen Styles

I. ESCAP issues

John Thompson distributed a document describing the issues that the Committee needs to resolve in order to make a recommendation about the future uses for the adjusted data. The document is attached. John walked the Committee through each of the concerns in the attachment.

John Thompson also asked the Committee to review the document distributed and to provide him with comments by close of business. The remainder of the meetings this week will be used to discuss these issues. Subsequent to the meeting, John incorporated the Committee's comments into the issues document. The revised document is also attached.

II. Relationship between census duplicates and erroneous enumerations

As stated above there are several issues remaining for the Committee to consider regarding duplicates. At the September 27th ESCAP meeting, John Thompson tasked Bob Fay with examining duplicates and their relationship with erroneous enumerations.

In order to understand the magnitude of duplication in the census, Bob Fay presented data illustrating what percent of duplication was identified by computer matching. This estimate, when used with the results of the work in computer matching, produced an estimate of the overall level of duplicate persons in the unadjusted census. Using evidence from the *Housing Unit Duplication Operations* and a reanalysis of the person duplication study, Bob Fay determined that computer matching identified approximately 75.7 percent of the duplicates. It was noted that this was a conservative estimate and thus, the level of duplication in the census was probably higher. Bob also presented a draft table illustrating the type of duplication, the A.C.E. measurement, and the issue for each type. Once this table is completed it will be presented to the Committee.

Bob Fay also presented data indicating a significant understatement of duplication in the A.C.E. These results are preliminary and more refined results will be presented on Wednesday, October 3.

III. Next Meeting

The next meeting is scheduled for October 3 at 9:00. The agenda is to discuss demographic analysis (DA), correlation bias, and measurements of erroneous enumerations and duplicates.

ESCAP MEETING NO. 75 - 10/01/01

HANDOUTS

ESCAP Issues

October 1, 2001

The previous ESCAP recommendation for the use of unadjusted Census 2000 results for redistricting identified issues associated with the A.C.E., demographic analysis, and Census 2000. In particular, differences between the A.C.E. and demographic analysis could not be explained leading to questions regarding both the accuracy of demographic analysis and the A.C.E. Concerns were also identified with balancing and the use of synthetic estimation.

Since March, the ESCAP has been supported by an extensive evaluation program designed to inform the issues described above. We now understand the causes and effects of balancing error. While we still have discrepancies between demographic analysis and the A.C.E., we have brought them into closer focus. Demographic analysis has incorporated Census 2000 data, and we have reexamined the assumptions that underlie the methodology. We have also identified those nonsampling errors that seem to have the greatest effects on the accuracy of the A.C.E. – accurate measurement of erroneous enumerations, and correlation bias. It appears that the A.C.E. overestimates the population at the national level due to problems in estimating the full level of Census 2000 erroneous enumerations, and demographic analysis may not include a full measure of the unauthorized population.

The issues we now face arise from the evaluation studies that have been conducted to measure the components of nonsampling error associated with the A.C.E. and demographic analysis. The evaluations while providing some measures, have a great deal of unresolved and conflicting information. Therefore we face issues in determining how much the A.C.E. is overstating the true population, and how much, if any, demographic analysis is understating the population. We also must associate levels of uncertainty with the evaluation studies, and with the models used to account for missing data and synthetic estimation in the A.C.E. In addition, the total error model methodology designed to produce an overall measure of the A.C.E. accuracy has been revised and new issues have been identified regarding whether it provides a complete accounting of all of the errors measured for the A.C.E.

The discussion below attempts to identify the issues we must resolve in order to reach a decision.

Demographic Analysis

Demographic analysis is still inconsistent with the A.C.E.

Demographic analysis has been reviewed and revised. Census 2000 data have been incorporated into the production of the new DA estimates as have the results of reviewing the underlying assumptions.

There are concerns that the level of unauthorized persons included in the DA estimates may be too low. Remaining issues are related to this concern. That is, are the DA estimates still too low because of under-estimation of unauthorized persons? It should be noted that if we are to explain the difference between DA and the A.C.E., we are assuming that the A.C.E. included a significant number of unauthorized persons not represented in the DA estimates.

We also have realized that we will gain insights into the differences between DA and the A.C.E. by comparing DA to A.C.E. results corrected for biases. This comparison may also provide validation for some of the bias measurements as well.

Measurement of Census 2000 Erroneous Enumerations

We have evidence that the A.C.E. did not measure all of the erroneous enumerations in Census 2000. We have an extremely wide range in which the actual measurement should be.

The key evaluation study – the Evaluation Followup (or Measurement Error Study) initially indicated that significant understatement of erroneous enumerations had occurred. However, the results were questioned and a review of the study was initiated. This review resulted in a large number of unresolved and conflicting results.

The measurement of erroneous enumerations is critical to both the national net undercount, and to subnational estimates. It should be noted that the case for adjustment could be strengthened by an accurate measurement of erroneous enumerations.

Resolving the questions related to measurement of erroneous enumerations is critical to our decision process. We may be able to examine these results in relation to demographic analysis, and potentially the duplication studies.

Measurement of Census 2000 Duplication

The level of Census 2000 duplication not included in the A.C.E. universe does not seem to be large enough to explain the differences between DA and the A.C.E. Furthermore, the duplication studies have indicated that the A.C.E. did not completely measure Census 2000 duplicates outside of the A.C.E. search area. Duplicates outside the A.C.E. search area should be included in the A.C.E. as other residence erroneous enumerations.

There are several issues associated with the duplication studies. First, they were conducted by computer matching only because we did not have the resources or time to match the A.C.E. E-sample to the entire country using both computer and clerical matching. Therefore, the computer matching understates the actual level of duplication that the A.C.E. should measure. When conducted in the A.C.E. blocks, the computer matching found about 37 percent of the actual duplicates.

The question that must be resolved is the level of duplicates that the computer matching picked up. Because these duplicates also represent a lower bound on other residence erroneous enumerations, a resolution of this issue could help explain the questions related to the overall measurement of erroneous enumerations.

Correlation Bias

Correlation bias is an important component of A.C.E. accuracy. Various assumptions regarding correlation bias have a large effect on the total error model measures of A.C.E. net accuracy. Correlation bias also leads to estimates of the total population that are inconsistent with demographic analysis. We have calculated a number of different estimates of correlation bias. We must consider the following:

Should we include correlation bias only for the Black population?
Should we assume correlation bias is the same for the Hispanic population as for the Black population?
Should we assume correlation bias to be the same for owners and renters?
We have various estimates of correlation bias ranging from 10 percent and up.
Should we require some consistency between the population estimates that result from correlation bias and demographic analysis?

Total Error Model

The total error model is designed to be an accounting of the errors in the A.C.E., and thus the basis for assessing the gains from the adjustment. We have developed a new total error model that is designed to incorporate the results of the current A.C.E. evaluations. There are several concerns and issues associated with the total error model:

Validation -- At this point we are not in a position to validate the accuracy of the total error model; that is, to determine whether it is correctly incorporating the evaluation findings.

Completeness – The total error model does not include all of the results from the balancing evaluations. In addition, the total error model does not treat the results from the matching error study and the evaluation followup consistent with the way in which they were treated in the previous model, and may be omitting additional error components.

We may have to rely on the individual studies, or go back to the 1990 model scaled to our new

findings.

Loss Function Analysis

The loss functions are based on the total error model. They assess the degree to which the adjusted and unadjusted data are closer to the target populations for a specified grouping of areas (e.g., counties within states). The target populations are based on a series of assumptions used in the total error model calculations such as the level of correlation bias etc. The target populations are also derived using synthetic estimation, leading to questions regarding the degree to which the loss function analysis is influenced by the synthetic component of the target populations. The weights used in the loss function analysis also imply particular importance of errors measured for various sized places. The loss function analysis is also conducted to assess numeric and distributive accuracy.

The issues we must address include the effect of the use of synthetic estimation in developing the target populations, the implications of the weighting, and the relative importance of numeric and distributive accuracy.

Balancing Error

We have resolved the issues associated with balancing error for the most part. We demonstrated that the primary cause of the discrepancy between the number of correct enumerations in blocks surrounding the A.C.E. sample and the number of P-sample matches was P-sample geocoding error. Since this has little effect on the accuracy of the A.C.E., most of our concerns are addressed. However, the Targeted Extended Search (TES) evaluations did identify A.C.E. error components that are not completely reflected in the total error model.

The remaining issue is the degree to which the total error model will overstate the accuracy of the A.C.E.

Missing Data

We have examined a variety of models to predict the effects of missing data. They give a fairly wide range under some assumptions. Given the materials we have examined, it appears that we have missing data models that both understate and overstate the effects of missing data on the A.C.E. estimates. We have chosen to represent these effects in the form of increased uncertainty in the A.C.E. estimates.

The issue remaining with missing data is whether we are over or under representing the full degree of uncertainty in the A.C.E. analyses.

Synthetic Estimation

Synthetic error affects both the adjusted and unadjusted census results. The error introduced by synthetic estimation is not included in the total error model, and cannot be estimated directly to assess the error in adjusting for undercount. We also know that the loss functions incorporate target populations derived using synthetic estimation. It is important to understand the potential effects of synthetic error, particularly on loss function analysis and therefore we use populations constructed from surrogate variables to simulate synthetic error.

The issues that must be considered are the degree to which the surrogate variables represent Census 2000 undercount, the degree to which the simulations of synthetic error are influenced by the construct of the surrogate variable populations, and the relative effect of the synthetic error on census and A.C.E. loss.

ESCAP MEETING NO. 75 - 10/01/01

HANDOUTS

ESCAP Issues Revised to Reflect ESCAP Comments

October 2, 2001

The previous ESCAP recommendation for the use of unadjusted Census 2000 results for redistricting identified issues associated with the A.C.E., demographic analysis, and Census 2000. In particular, differences between the A.C.E. and demographic analysis could not be explained, leading to questions regarding both the accuracy of demographic analysis and the A.C.E. Concerns were also identified with balancing and the use of synthetic estimation.

Since March, the ESCAP has been supported by an extensive evaluation program designed to inform the issues described above. We now understand the causes and effects of balancing error. While we still have discrepancies between demographic analysis and the A.C.E., we have brought them into closer focus. Demographic analysis has incorporated Census 2000 data, and we have reexamined the assumptions that underlie the methodology. We have also identified those nonsampling errors that seem to have the greatest effects on the accuracy of the A.C.E. – accurate measurement of erroneous enumerations, and correlation bias. It appears that the A.C.E. overestimates the population at the national level due to problems in estimating the full level of Census 2000 erroneous enumerations, and demographic analysis may not include a full measure of the unauthorized population.

The issues we now face arise from the evaluation studies that have been conducted to measure the components of nonsampling error associated with the A.C.E. and demographic analysis. The evaluations while providing some measures, have a great deal of unresolved and conflicting information. Therefore we face issues in determining how much the A.C.E. is overstating the true population, and how much, if any, demographic analysis is understating the population. We also must associate levels of uncertainty with the evaluation studies, and with the models used to account for missing data and synthetic estimation in the A.C.E. In addition, the total error model methodology designed to produce an overall measure of A.C.E. accuracy has been revised and new issues have been identified regarding whether it provides a complete accounting of all of the errors measured for the A.C.E. For the recommendation in March, we had to rely on 1990 PES evaluations for most of the total error model components, and that was a significant limitation. Now we have data from the 2000 evaluations. However, we need a method of assessing their combined effects, and the current total error model does not appear to be including all components of the 2000 evaluations.

The discussion below attempts to identify the issues we must resolve in order to reach a decision.

Demographic Analysis

Demographic analysis is still inconsistent with the A.C.E.

Demographic analysis has been reviewed and revised. Census 2000 data have been incorporated into the production of the new DA estimates as have the results of reviewing the underlying assumptions.

There are concerns that the level of unauthorized persons included in the DA estimates may be too low. Remaining issues are related to this concern. That is, are the DA estimates still too low because of under-estimation of unauthorized persons. It should be noted that if we are to explain the difference between DA and the A.C.E., we are assuming that the A.C.E. included a significant number of unauthorized persons not represented in the DA estimates.

There are concerns that some of the components of the foreign-born are not well measured in demographic analysis. This uncertainty could lead to the DA estimate being either too high or too low. If the estimates are too low, then some of the difference between ACE and DA might be explainable.

We should explain how DA estimates the "residual immigrants" and how the pieces of this process are subject to uncertainty. It is important because it presumably is the biggest source of uncertainty in the DA estimates (though uncertainty about birth registration completeness and emigration still isn't trivial). We should say something that indicates, in a concrete way, why there is uncertainty about the DA estimates.

Measurement of Census 2000 Erroneous Enumerations

We have evidence that the A.C.E. did not measure all of the erroneous enumerations in Census 2000. We have an extremely wide range in which the actual measurement should be.

The key evaluation study – the Evaluation Followup (or Measurement Error Study) – initially indicated that significant understatement of erroneous enumerations had occurred. However, the results were questioned and a review of the study was initiated. This review resulted in a large number of unresolved and conflicting results. It is important to understand the unresolved and conflicting cases.

The measurement of erroneous enumerations is critical to both the national net undercount, and to subnational estimates. It should be noted that the case for adjustment could be strengthened by an accurate measurement of erroneous enumerations.

Resolving the questions related to measurement of erroneous enumerations is critical to our decision process. We may be able to examine these results in relation to demographic analysis, and potentially the duplication studies.

Measurement of Census 2000 Duplication

The level of Census 2000 duplication not included in the A.C.E. universe does not seem to be large enough to explain the differences between DA and the A.C.E. The A.C.E. was not designed to estimate census duplicates between housing units and GQs versus within the housing unit population. However, a critical issue is that the duplication studies have indicated that the A.C.E. did not completely measure Census 2000 duplicates outside of the A.C.E. search area. Duplicates outside the A.C.E. search area should be included in the A.C.E. as other residence erroneous enumerations.

There are several issues associated with the duplication studies. First, they were conducted by computer matching only because we did not have the resources or time to match the A.C.E. E-sample to the entire country using both computer and clerical matching. Therefore, the computer matching understates the actual level of duplication that the A.C.E. should measure. When conducted in the A.C.E. blocks, the computer matching found about 37 percent of the actual duplicates.

The question that must be resolved is the level of duplicates that the computer matching picked up. Because these duplicates also represent a lower bound on other residence erroneous enumerations, a resolution of this issue could help explain the questions related to the overall measurement of erroneous enumerations.

For the total error model and loss functions, we need the evaluations to provide an accurate (or at least approximately unbiased) estimate of errors in the A.C.E. estimates, including errors related to this underestimation of census duplicates. This is important because if the A.C.E. underestimates duplication and this is not measured in the evaluations, then we could make an incorrect decision.

Correlation Bias

Correlation bias is an important component of the A.C.E. accuracy. Various assumptions regarding correlation bias have a large effect on the total error model measures of A.C.E. net accuracy. Correlation bias also leads to estimates of the total population that are inconsistent with demographic analysis. This argument is a bit circular since DA sex ratios are used to estimate correlation bias. We have calculated a number of different estimates of correlation bias. We must consider the following:

Should we include correlation bias only for the Black population?
Should we assume correlation bias is the same for the Hispanic population as for the Black population?
Should we assume correlation bias to be the same for owners and renters?
We have various estimates of correlation bias ranging from 10 percent and up.
Should we require some consistency between the population estimates that result from correlation bias and demographic analysis?

Total Error Model

The total error model is designed to be an accounting of the errors in the A.C.E., and thus the basis for assessing the gains from the adjustment. We have developed a new total error model that is designed to incorporate the results of the current A.C.E. evaluations. We have addressed individual errors through the A.C.E. evaluations, and it is critical to have a total error model that accurately combines the results of the evaluations. There are several concerns and issues associated with the total error model:

Validation -- At this point; we are not in a position to validate the accuracy of the total error model; that is, to determine whether it is correctly incorporating the evaluation findings.

Completeness – The total error model does not include all of the results from the balancing evaluations. In addition, the total error model does not treat the results from the matching error study and the evaluation followup consistent with the way in which they were treated in the previous model, and may be omitting additional error components.

We may have to rely on the individual studies, or go back to the 1990 model scaled to our new findings.

Loss Function Analysis

The loss functions are based on the total error model. They assess the degree to which the adjusted and unadjusted data are closer to the target populations for a specified grouping of areas (e.g., counties within states). The target populations are based on a series of assumptions used in the total error model calculations such as the level of correlation bias etc. The target populations are also derived using synthetic estimation, leading to questions regarding the degree to which the loss function analysis is influenced by the synthetic component of the target populations. The weights used in the loss function analysis also imply particular importance of errors measured for various sized places. The loss function analysis is also conducted to assess numeric and distributive accuracy.

The issues we must address include the effect of the use of synthetic estimation in developing the target populations, the implications of the weighting, and the relative importance of numeric and distributive accuracy.

Balancing Error

We have resolved the issues associated with balancing error for the most part. We demonstrated that the primary cause of the discrepancy between the number of correct enumerations in blocks surrounding the A.C.E. sample and the number of P-sample matches was P-sample geocoding error. Since this has little effect on the accuracy of the A.C.E., most of our concerns are addressed. However, the Targeted Extended Search (TES) evaluations did identify A.C.E. error components that are not completely reflected in the total error model.

The remaining issue is the degree to which the total error model will overstate the accuracy of the A.C.E.

Missing Data

We have examined a variety of models to predict the effects of missing data. They give a fairly wide range under some assumptions. Given the materials we have examined, it appears that we have missing data models that both understate and overstate the effects of missing data on the A.C.E. estimates. We have chosen to represent these effects in the form of increased uncertainty in the A.C.E. estimates.

The issue remaining with missing data is whether we are over or under representing the full degree of uncertainty in the A.C.E. analyses.

Synthetic Estimation

Synthetic error affects both the adjusted and unadjusted census results. The error introduced by synthetic estimation is not included in the total error model, and cannot be estimated directly to assess the error in adjusting for undercount. We also know that the loss functions incorporate target populations derived using synthetic estimation. It is important to understand the potential effects of synthetic error, particularly on loss function analysis, and therefore we use populations constructed from surrogate variables to simulate synthetic error.

The issues that must be considered are the degree to which the surrogate variables represent Census 2000 undercount, the degree to which the simulations of synthetic error are influenced by the construct of the surrogate variable populations, and the relative effect of the synthetic error on census and A.C.E. loss.

ESCAP MEETING NO. 76 - 10/03/01

AGENDA

Kathleen P Porter
09/20/2001 02:04 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, RJones17@doc.gov, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, roxie.jones@mail.doc.gov, RJones17@doc.gov, Jacqueline M Cusick/DIR/HQ/BOC@BOC
cc: Bonnie J Demarr/DSSD/HQ/BOC@BOC, Elizabeth A Krejsa/PRED/HQ/BOC@BOC, Tamara S Adams/DSSD/HQ/BOC@BOC, Mary Helen Mulry/SRD/HQ/BOC@BOC, Roxanne Feldpausch/DSSD/HQ/BOC@BOC
Subject: ESCAP Meetings for week of September 24

The ESCAP Meetings for the week of September 24 are as follows (all are in Rm. 2412/3):

September 25 10:30-12:00 Correlation Bias - Bell (SRD)

September 26 10:30-12:00 Evaluation and Person Followup Questionnaires - Martin (DIR)
EFU Rework - Adams/Krejsa (DSSD/PRED)
Person Dups and EEs - Feldpausch (DSSD)

September 28 1:30-3:00 Final Total Error and Loss Functions - Mulry
[**Note**: the topic for the September 28 meeting was changed to "Alternative DA Scenarios." A revised agenda for meeting #74 was not distributed.]

ESCAP MEETING NO. 76 - 10/03/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting #76

October 3, 2001

Prepared by: Nick Birnbaum

The seventy-sixth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on October 3, 2001 at 9:00 am. The agenda for the meeting was to: 1) discuss alternative assumptions for the foreign-born population, 2) review correlation bias briefly, and 3) follow up on the discussion from the meeting of October 1, 2001 regarding the examination of duplicates and their relationship to erroneous enumerations.

Committee Attendees:

John Thompson
Ruth Ann Killion
Cynthia Clark
Jay Waite
Bob Fay
Howard Hogan
John Long
Teresa Angueira
Nancy Potok
Nancy Gordon
Carol Van Horn

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines	Bill Bell
Kevin Deardorff	Nick Birnbaum
Sarah Brady	Maria Urrutia
Carolee Bush	Raj Singh

I. Alternative Assumptions for the Foreign-Born Population

Population Division staff presented data providing a range of estimates for the foreign-born population, including the revised demographic analysis (DA) estimate, which assumes a 15% undercount for the residual foreign born (the known components plus the (implied) unauthorized). Population Division also constructed three other estimates for the foreign-born population – two higher than the revised DA estimate and one lower. Only the 15% revised DA estimate has the full complement of detailed data disaggregated by age, race (Black/non-Black), and sex.

The purpose of the analysis was to show the implications of different estimates of the foreign-born population and its components. For example, the DA estimate of the foreign-born population that is the highest, results in a total population estimate that approaches the A.C.E. population level. However, this estimate assumes a 20% undercount of implied unauthorized migrants and would require that this component constitute 28% of the foreign-born population. The revised DA estimate assumes a 15% undercount of the residual foreign born (both the known components and implied unauthorized migrants) and implies that unauthorized migrants represent 25% of the foreign-born population.

Another way to gauge the plausibility of the various estimates of the foreign-born population is to look at the implied growth ratios of legal immigrants to unauthorized migrants during the 1990s. Given that the growth to legal immigrants during the 1990s was 7.5 million (according to INS data), this implies a growth ratio of legal immigrants to unauthorized migrants of only 1.4 to 1 for the highest estimate of the foreign-born population.

Because virtually all of the variability in the DA estimate of the total population is from the foreign-born components, one can examine the plausibility of the various estimates of the foreign born relative to the corresponding DA estimates of the total population. For example, the Census 2000 level estimate of the foreign born indicates that this group constitutes 11.1% of the total population. Under the revised DA estimate, the foreign-born population represents 11.6% of the total. Under the highest estimate of the foreign born, they would be 12.1% of the population, fully one percentage point above the Census 2000 estimate of their proportion of the total population. This estimate implies a *net* undercount, relative to Census 2000, of almost 2 million.

Population Division staff discussed the range of estimates presented and noted that it was a fairly extreme range. John Thompson then requested that the staff determine a plausible range for the DA estimate of the total population.

II. Review of Correlation Bias

John Thompson briefly summarized the work on correlation bias being conducted for the impending decision. Of the various DA estimates presented to the Committee, only the revised DA estimate (assumes 15% undercount of the residual foreign born) will provide sufficiently detailed data for the estimation of correlation bias. However, once the level of correlation bias is estimated for one or more demographic groups, different proportions of these levels (for example, 25%, 50%, 75%, etc.) will be used in the total error model and loss function analyses to determine how sensitive the results are to varying levels of correlation bias.

III. Examination of Duplicates and Their Relationship to Erroneous Enumerations

As a follow-up to the discussion at the meeting of October 1, 2001, Bob Fay presented additional data and analyses related to this issue. The data indicate, by various types of duplication cases, the potential for a non-trivial level of errors not accounted for in the total error model. Bob Fay will continue his research to determine the extent to which these A.C.E. errors are not included in the total error model.

IV. Next Meeting

The agenda for the next meeting, scheduled for October 4, 2001, is to revisit the total error model.

ESCAP MEETING NO. 77 - 10/04/01

AGENDA

Kathleen P Porter
09/20/2001 02:04 PM

To: Barbara E Hotchkiss/DSD/HQ/BOC@BOC, Betty Ann Saucier/DIR/HQ/BOC@BOC, Carnelle E Sligh/PRED/HQ/BOC@BOC, Carol A Campbell/DMD/HQ/BOC@BOC, Carol M Van Horn/DIR/HQ/BOC@BOC, Carolee Bush/DIR/HQ/BOC@BOC, Cecilia R Lewis/DMD/HQ/BOC@BOC, Charles T Lee Jr/DMD/HQ/BOC@BOC, Cynthia Z F Clark/DIR/HQ/BOC@BOC, Deborah A Fenstermaker/DSSD/HQ/BOC@BOC, Deena Grover/DSSD/HQ/BOC@BOC, Donna L Kostanich/DSSD/HQ/BOC@BOC, Fay F Nash/DMD/HQ/BOC@BOC, Hazel V Beaton/SRD/HQ/BOC@BOC, Howard R Hogan/DSSD/HQ/BOC@BOC, John F Long/POP/HQ/BOC@BOC, John H Thompson/DIR/HQ/BOC@BOC, Kathleen M Styles/DMD/HQ/BOC@BOC, Linda A Hiner/DSSD/HQ/BOC@BOC, Lois M Kline/POP/HQ/BOC@BOC, Margaret A Applekamp/DIR/HQ/BOC@BOC, Maria E Urrutia/DMD/HQ/BOC@BOC, Marvin D Raines/DIR/HQ/BOC@BOC, Mary A Cochran/DIR/HQ/BOC@BOC, Mary E Williams/DIR/HQ/BOC@BOC, Nancy A Potok/DIR/HQ/BOC@BOC, Nancy M Gordon/DSD/HQ/BOC@BOC, Nicholas I Birnbaum/DMD/HQ/BOC@BOC, Patricia E Curran/DIR/HQ/BOC@BOC, Phyllis A Bonnette/DIR/HQ/BOC@BOC, Preston J Waite/DMD/HQ/BOC@BOC, Rajendra P Singh/DSSD/HQ/BOC@BOC, Rita J Petroni/PRED/HQ/BOC@BOC, Robert E Fay III/DIR/HQ/BOC@BOC, RJones17@doc.gov, Ruth Ann Killion/PRED/HQ/BOC@BOC, Sarah E Brady/DMD/HQ/BOC@BOC, Sue A Kent/DMD/HQ/BOC@BOC, Teresa Angueira/DMD/HQ/BOC@BOC, Tommy Wright/SRD/HQ/BOC@BOC, William G Barron Jr/DIR/HQ/BOC@BOC, William R Bell/SRD/HQ/BOC@BOC, roxie.jones@mail.doc.gov, RJones17@doc.gov, Jacqueline M Cusick/DIR/HQ/BOC@BOC
cc: Bonnie J Demarr/DSSD/HQ/BOC@BOC, Elizabeth A Krejsa/PRED/HQ/BOC@BOC, Tamara S Adams/DSSD/HQ/BOC@BOC, Mary Helen Mulry/SRD/HQ/BOC@BOC, Roxanne Feldpausch/DSSD/HQ/BOC@BOC
Subject: ESCAP Meetings for week of September 24

The ESCAP Meetings for the week of September 24 are as follows (all are in Rm. 2412/3):

September 25 10:30-12:00 Correlation Bias - Bell (SRD)

September 26 10:30-12:00 Evaluation and Person Followup Questionnaires - Martin (DIR)
EFU Rework - Adams/Krejsa (DSSD/PRED)
Person Dups and EEs - Feldpausch (DSSD)

September 28 1:30-3:00 Final Total Error and Loss Functions - Mulry

ESCAP MEETING NO. 77 - 10/04/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting # 77

October 4, 2001

Prepared by: Sarah Brady

The seventy-seventh meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on October 4, 2001 at 11:00. The agenda for the meeting was to discuss issues regarding what is included in the total error model.

Committee Attendees:

Nancy Potok
John Thompson
Nancy Gordon
Cynthia Clark
Jay Waite
Carol Van Horn
Bob Fay
Teresa Angueira
Howard Hogan
Ruth Ann Killion
John Long

Deputy Director/Acting Director:
William Barron

Other Attendees:

Bill Bell
Marvin Raines
Raj Singh
Maria Urrutia
Sarah Brady
Kathleen Styles
Rita Petroni

I. Total error model– What is included

John Thompson discussed the fact that the total error model does not include all the components of error in the A.C.E. as measured by the evaluations. John handed out tables illustrating what components of error were included, partially included, and not included at all. Work must be done to incorporate these components correctly into the total error model. John indicated that he, Jay Waite, and staff would proceed in working through the weekend to discuss these issues (including Bob Fay's work that was described in the previous minutes). John expects the results to be available to the Committee for Tuesday's meeting.

II. Next Meeting

The next meeting is scheduled for October 9. The agenda is to discuss the remaining unresolved issues.

ESCAP MEETING NO. 78 - 10/09/01
AGENDA

"There was no agenda developed or used for the October 9, 2001 meeting."

ESCAP MEETING NO. 78 - 10/09/01

MINUTES

Minutes of the Executive Steering Committee on
Accuracy and Coverage Evaluation (A.C.E.) Policy (ESCAP) Meeting #78

October 9, 2001

Prepared by: Nick Birnbaum

The seventy-eighth meeting of the Executive Steering Committee on Accuracy and Coverage Evaluation Policy was held on October 9, 2001 at 10:00 am. The agenda for the meeting was to discuss the current status of the Committee's work in preparation for the issuance of its recommendation.

Committee Attendees:

John Thompson
Ruth Ann Killion
Cynthia Clark
Jay Waite
Bob Fay
Howard Hogan
John Long
Teresa Angueira
Nancy Potok
Nancy Gordon
Carol Van Horn

Deputy Director/Acting Director:
William Barron

Other Attendees:

Marvin Raines	Bill Bell
Kathleen Styles	Nick Birnbaum
Sarah Brady	Maria Urrutia
Donna Kostanich	Raj Singh
Rita Petroni	

I. Review of Current Status of the Committee's Work

John Thompson distributed a draft document summarizing preliminary conclusions and outstanding issues relating to the Committee's work. This document is attached. John briefly discussed each item as he walked the Committee through the document. Based on clarifications obtained during the Committee discussion, the document was subsequently revised. The revised version is also attached.

John also spoke briefly about the work done over the weekend to attempt to resolve the issues with the measurement of A.C.E. errors and with the total error model (TEM). (See items 6 through 8 in the handout.) John then turned to other Committee members for their comments on the handout and his assessment of these issues. It was noted that the Committee was going to have to base its recommendation on the results of the studies presented earlier, since the total error model was not going to be available for some time.

Bob Fay discussed his additional work relating to the identification of duplicates by the A.C.E. (See reference to duplicate studies in item 6.2 in the handout.) This work indicates a very serious problem with the A.C.E. measurement of duplicate enumeration.

At this point, John called for the Committee to enter into deliberations and only Committee members and senior technical staff remained for this segment of the meeting.

II. Next Meeting

Any additional ESCAP meetings to be held prior to the issuance of the Committee's recommendation will be a continuation of the Committee's deliberations, and attendance will be limited to members and senior technical staff as noted above.

ESCAP MEETING NO. 78 - 10/09/01

HANDOUTS

Status of ESCAP2 Review
Draft October 9, 2001

1. Demographic analysis (DA) and the A.C.E. indicate a differential undercount for the Black population. Demographic analysis is consistent with the A.C.E. in estimating an undercount for the Black population. Demographic analysis also estimates a higher undercount for Black males giving support to expectations of some level of correlation bias.

2. Demographic analysis and the A.C.E. are inconsistent for the non-Black population. The A.C.E. indicates an undercount, and DA indicates an overcount. For the non-Black population, DA is about 3 million lower than the A.C.E., while for the Black population DA is about 300,000 higher.

3. Both DA and the A.C.E. have been extensively evaluated. We have reworked DA in consultation with outside experts, and carefully examined its main area of uncertainty, the foreign-born population. We have also conducted a number of studies of the A.C.E. The main sources of information on accuracy have come from the Matching Error Study (MES), the Measurement Error Review (MER) or Evaluation Followup (EFU), Duplicate Review, Targeted Extended Search (TES), review of missing data alternatives, and Synthetic Error studies.

4. The A.C.E. evaluations have removed concerns about balancing, conditioning, and Census 2000 reinstated and imputed persons.

5. The A.C.E. evaluations give some evidence that the A.C.E. does not include a complete measure of Census 2000 erroneous enumerations, and that taking all errors into account, the A.C.E. is overestimating the population at a **net level**. This is consistent in a very general sense with DA, however we are uncertain as to the level of A.C.E. overstatement.

6. The major inputs to the total error model are the MES, EFU, uncertainty measures for missing data, correlation bias derived from DA, and sampling error. It is not clear whether these studies capture all of the error associated with the A.C.E., or whether all components of these errors are being incorporated correctly into the total error model.

6.1 The total error model cannot be validated at this point. It is not clear whether it can be validated before October 15.

6.2 The MER/EFU provides estimates of net error for both the P and E samples. Results for the E-sample contain a number of unresolved and conflicting cases. Depending on the assumptions made regarding conflicting and unresolved cases, the net bias measured for the E-sample ranges from 1.5 million to about 2.9 million. The duplicate studies indicate that these measures potentially do not include an additional 1 million errors.

6.3 The MER/EFU results for the P-sample indicate that the net bias is about -440,000. This result is the net effects of EFU measures of fairly large changes in mover status. That is, nonmovers becoming movers, and movers becoming nonmovers. It appears that the EFU may be mis-classifying some movers as non movers due to design features of the questionnaire. If this is the case, an argument could be made that the -440,000 should in fact be positive.

6.4 The MES indicates that the net error due to matching is about 480,000. This is mostly due to errors in matching the P-sample, with roughly only 40,000 errors in the E-sample.

6.5 There is some overlap between the MES and MER/EFU studies. Part of the issues with the total error model relates to determining the degree of this overlap.

6.6 Correlation bias accounts for a net bias of about -750,000 to -1.3 million depending on which model for correlation bias is assumed. For loss function analysis we have also examined ranges of correlation bias from 0 to 100 percent within each of the models.

7. Synthetic error remains somewhat of a mystery because the revised analysis depended on

the 1990 measures of error. Synthetic error must be considered when the results of loss function analysis are reviewed.

8. Missing data can have a fairly large effect on the dual system estimates under certain non-ignorable missing data models. We had decided to include these effects in the total error model as a random effect. There is also a new bias term in the missing data model which appears to now have a significant effect. More discussion of this is probably warranted when we have validated the total error model.

ESCAP MEETING NO. 78 - 10/09/01

HANDOUTS

Status of ESCAP2 Review
Draft (**Revision to October 9 version**)

1. Demographic analysis (DA) and the A.C.E. indicate a differential undercount for the Black population. Demographic analysis is consistent with the A.C.E. in estimating an undercount for the Black population. Demographic analysis also estimates a higher undercount for Black males giving support to expectations of some level of correlation bias.

2. Demographic analysis and the A.C.E. are inconsistent for the non-Black population. The A.C.E. indicates an undercount, and DA indicates an overcount. For the non-Black population, DA is about 3 million lower than the A.C.E., while for the Black population DA is about 300,000 higher.

3. Both DA and the A.C.E. have been extensively evaluated. We have reworked DA in consultation with outside experts, and carefully examined its main area of uncertainty, the foreign-born population. We have also conducted a number of studies of the A.C.E. The main sources of information on accuracy have come from the Matching Error Study (MES), the Measurement Error Review (MER) or Evaluation Followup (EFU), Duplicate Review, Targeted Extended Search (TES), review of missing data alternatives, and Synthetic Error studies.

4. The A.C.E. evaluations have removed concerns about balancing, conditioning, and Census 2000 reinstated and imputed persons.

5. The A.C.E. evaluations give evidence that the A.C.E. does not include a complete measure of Census 2000 erroneous enumerations, and that taking all errors into account, the A.C.E. is significantly overestimating the population at a **net level**. This is consistent in a very general sense with DA, however we are uncertain as to the level of A.C.E. overstatement.

6. The major inputs to the total error model are the MES, EFU, uncertainty measures for missing data, correlation bias derived from DA, and sampling error. It is clear that these studies do not capture all of the error associated with the A.C.E. and that all components of these errors are not being incorporated correctly into the total error model.

6.1 The total error model cannot be validated at this point. It is not clear whether it can be validated before October 15.

6.2 The MER/EFU provides estimates of net error for both the P and E samples. Results for the E-sample contain a number of unresolved and conflicting cases. Depending on the assumptions made regarding conflicting and unresolved cases, the net bias measured for the E-sample ranges from 1.5 million to over 2.9 million. The duplicate studies indicate that these measures potentially do not include an additional 1 million or more errors.

6.3 The MER/EFU results for the P-sample indicate that the net bias is about -440,000. This result is the net effects of EFU measures of fairly large changes in mover status. That is, nonmovers becoming movers, and movers becoming nonmovers. It appears that the EFU may be mis-classifying some movers as nonmovers due to design features of the questionnaire. If this is the case, an argument could be made that the -440,000 should in fact be positive.

6.4 The MES indicates that the net error due to matching is about 480,000. This is mostly due to errors in matching the P-sample, with only about 40,000 errors in the E-sample.

6.5 There is some overlap between the MES and MER/EFU studies. Part of the issues with the total error model relates to determining the degree of this overlap.

6.6 Correlation bias accounts for a net bias of about -750,000 to -1.3 million depending on which model for correlation bias is assumed. For loss function analysis we have also examined ranges of correlation bias from 0 to 100 percent within each of the models.

7. Synthetic error remains somewhat of a mystery because the revised analysis depended on the 1990 measures of error. Synthetic error must be considered when the results of loss function analysis are reviewed.

8. Missing data can have a fairly large effect on the dual system estimates under certain non-ignorable missing data models. We had decided to include these effects in the total error model as a random effect. There is also a new bias term due to missing data in the total error model. More discussion of this is warranted when we have validated the total error model.